Spectrotemporal Organization of Excitatory and Inhibitory Receptive Fields of Cat Posterior Auditory Field Neurons

WILLIAM C. LOFTUS AND MITCHELL L. SUTTER
Center for Neuroscience and Section of Neurobiology, Physiology and Behavior, University of California, Davis, California 95616

Received 28 August 2000; accepted in final form 27 February 2001

INTRODUCTION

In contrast to our understanding of the functional organization of different nonprimary auditory cortical fields in echolocating mustached bats, the response properties of neurons outside of the primary auditory cortex (A1) in nonecholocating mammals remains poorly understood (for reviews, see Ehret 1997; Suga 1988). A nonprimary auditory cortical area that is the subject of increasing interest is the posterior auditory field (PAF, a.k.a. field “P”) (Morel and Imig 1987) of cats. Earlier reports indicate that neurons in PAF, as in A1, respond to tone onsets and are tonotopically organized (Phillips and Orman 1984; Reale and Imig 1980). But there are numerous differences between A1 and PAF. In contrast to A1, PAF neurons exhibit longer response latencies, have “late” discharges following the initial onset response, have lower temporal precision, and are often tuned for intensity (Kitzes and Hollrigel 1996; Phillips and Orman 1984; Phillips et al. 1995, 1996). The sources of thalamic input to PAF and A1 and intracortical connections support physiological differences between the fields (Imig and Morel 1983; Kitzes and Hollrigel 1996; Morel and Imig 1987; Rodrigues-Dagaeff et al. 1989).

Despite these advances, our understanding of fundamental PAF receptive field (RF) properties is both scant and equivocal. Studies of intensity tuning have been mainly confined to characteristic frequency (CF), the stimulus frequency that can evoke a response at the lowest intensity. The earliest quantitative study of PAF response properties beyond CF (Phillips and Orman 1984) reported narrow, v-shaped, or circumscribed excitatory frequency tuning curves (eFTCs), comparable to A1. This is somewhat contradicted by a more recent investigation (Heil and Irvine 1998b) reporting a significant proportion of PAF neurons with broad, multipeaked, or patchy excitatory eFRAs. The temporal properties of PAF responses have also been the subject of limited study. In many PAF neurons, pure-tone stimuli elicit late responses in addition to early onset responses (Heil and Irvine 1998b; Phillips and Orman 1984). Real and Imig (1980) observed that late responses of neuron clusters in auditory cortex are generally more broadly tuned than the early responses. Yet investigations have focused on the onset responses of single neurons in PAF and have not examined the late responses. One goal of this study was to test the notion that the more complex RF properties reported in recent studies might relate to late discharges by describing how the timing of individual spikes relate to eFRA structure.

In describing PAF neurons, previous studies have focused on excitatory, rather than inhibitory, tuning curves and FRAs of PAF neurons. The contribution of inhibition to the shaping of eFRAs in PAF is supported by two considerations: there is an abundance of intensity-tuned neurons in PAF and inhibition...
overlapping with CF excitation is necessary for the emergence of intensity tuning. While several intriguing hypotheses have been advanced with regard to how inhibition shapes PAF response properties (Heil and Irvine 1998b; Phillips 1988), direct measurements of inhibitory FRAs (iFRAs), to our knowledge, have not been performed in PAF. Therefore another goal of this study was to provide the first direct measurements of iFRAs in PAF.

In this manuscript we characterize the spectral and temporal organization of excitatory frequency selectivity in PAF and the inhibitory receptive field structure outside the classical pure-tone excitatory FRA. This will lead to a more comprehensive view of the RF structure of PAF neurons. Parts of this investigation were presented in abstract form (Loftus and Sutter 1999).

**Methods**

**Surgical preparation**

This report is based on single-unit recordings from six left hemispheres and two right hemispheres in eight young adult cats, free of outer and middle ear infections. Anesthesia was induced with ketamine HCl (10 mg/kg im) and acetylpromazine maleate (0.28 mg/kg). To maintain areflexia during surgery and recording (≤96 h), pentobarbital sodium (Nembutal) in physiological saline solution was administered intravenously (2 mg·kg⁻¹·h⁻¹) with supplemental injections as needed. Lactated Ringer solution in 5% dextrose was administered through a separate catheter for a total volume of fluids of 4 ml·kg⁻¹·h⁻¹. Dexamethose sodium phosphate (0.14 mg/kg sc) and atropine sulfate (1 mg sc) were given to prevent brain edema and reduce mucous secretion, respectively. A tracheal cannula was inserted, and breathing was unassisted. Rectal temperature was monitored and held near 38°C by a heated water blanket.

The head was fixed by palato-orbital restraint, leaving the external meat unobstructed. We retracted the temporal muscle in one hemisphere, made an ~1 cm wide opening in the skull over the posterior ectosylvian sulcus, and reflected the dura providing access to PAF and portions of A1. Enlargements of the original opening, if necessary, were performed with rongeurs. The cortex was covered with silicone oil and an image of the surface vasculature was captured to mark the sites of electrode penetration.

**Acoustic stimulation**

The cat was placed in a double-walled sound-shielded booth (IAC). Sounds were presented in the near field via calibrated speakers (Radio Shack Optimus Pro-7AV and Radio Shack dual radial horn tweeter, No. 40-1377, with a crossover circuit at 7 kHz) driven by an amplifier (Radio Shack MPA-200). Speakers were at ±90° azimuth and 0° elevation relative to the animal, oriented directly toward the pinna contralateral to the recorded hemisphere, and 3.0 ft from the center of the head. The sound system was calibrated with a sound-level meter (Brüel and Kjaer type 2231) with a probe microphone positioned near the pinna. The frequency response of the system was essentially flat from 0.5 to 40 kHz except for two notches of <14 dB (peak to trough) centered at 1.2 and 2.8 kHz; otherwise, major resonances deviated less than ±6 dB from the average level (see discussion for possible influences of these resonances). Above 40 kHz, the output rolled off at a rate of 37 dB/octave. Harmonic distortion was <60 dB below the primary.

Stimuli were produced digitally (Tucker Davis Technology) with a 16-bit D/A converter at 120 kHz, attenuated, and low-pass filtered (TDT FT5 antialiasing filter;Fc = 50 kHz). A passive attenuator (Leader Model LAT-45) provided additional attenuation. Tone bursts were 50 ms with a 3-ms rise/fall time.

**Recording procedures**

Cortical activity was recorded extracranially with Parylene-coated tungsten microelectrodes (impedances 1–10 MΩ at 1 kHz). Electrodes were advanced into the caudal bank of the posterior ectosylvian sulcus, ~200–1,200 μm from the sulcus and ~3,300 μm deep with a hydraulic microdrive. In each experiment, a few initial penetrations were made in A1 to verify normal A1 responsiveness. The borders of PAF were defined using Reale and Imig’s (1980) criteria, including PAF’s reversal of tonotopy on the borders with A1 and VPAF. Activity from the electrode was amplified (AM Systems, model 1800), band-pass filtered, passed through a window discriminator (Bak Electronics, model RP1), and monitored on an oscilloscope and audio monitor. Search stimuli were tones of varying frequency and intensity. Once a neuron was isolated, stimuli were presented systematically while spike counts were monitored graphically. Each spike’s time of occurrence relative to stimulus onset (spike-time) was saved to disk. Spike-times were not adjusted for the travel time between the speakers and the tympanic membrane (calculated to be 2.65 ms). The recording window was 100 or 200 ms. In later experiments, spike waveforms were saved to disk and single units were isolated on and off-line in software.

**FRAs**

FRA collection was nearly identical to that in Sutter et al. (1999). Pure-tone FRAs were generated by presenting tones in a pseudorandom sequence of different frequency/intensity combinations (inter-stimulus interval: 400–2,000 ms). The stimulus presentation sequence was displayed graphically to confirm that it did not contain sequential clustering of presentations within the frequency-intensity space. Usually 45 frequencies (logarithmically spaced) and 15 intensities (over a 75-dB range) were used, providing 675 different frequency/intensity combinations per FRA. The majority of FRAs were recorded over a frequency range of 4 octaves, although the tested frequency range varied between 2 and 5.64 octaves, depending on the *estimated* range of inhibitory and excitatory responses. FRA graphs were not corrected for the frequency response of the sound-generation system; however, measures of response threshold and FRA classification were. In particular, spectral notches in the transfer function did not account for any regions of weak responsiveness separating the response peaks in multipeaked or patchy-unclassifiable (definitions in RESULTS) eFRAs. After a pure-tone FRA was obtained, a two-tone FRA(s) was collected to characterize the neuron’s two-tone inhibitory (TTI) response areas. This procedure, described in Sutter et al. (1999), used a two-tone simultaneous masking paradigm. Stimuli consisted of two simultaneous gated pure tones (inter-stimulus interval: 900–2,000 ms); a tone at the neuron’s best excitatory frequency (BEF) and a “variable” tone. The BEF and varying tone were combined digitally, and the resulting complex stimulus was played from one channel. The BEF tone was presented near the minimum intensity required to drive the neuron repetitively (usually 10–20 dB above threshold). The variable tones’ frequencies and intensities varied pseudo-randomly over successive presentations. When possible, the selected range of frequencies and intensities of the variable tones were identical to that of the pure-tone FRA so that the coordinate systems of the two-tone and the pure-tone FRAs could be superimposed. Sometimes, though, we had to extend the frequency range of two-tone FRAs to capture the full extent of inhibitory bands.

The purpose of the varying tone was to determine the FRA frequency/intensity regions where reduction of BEF tone responses occurred. Previous simultaneous two-tone masking studies have demonstrated that much of TTI observed in midbrain, thalamus, and cortex cannot be explained by basilar membrane two-tone suppression and thus have adopted the term inhibition (e.g., Fuzessery and Feng 1982; Imig et al. 1997; Suga and Tsuzuki 1985; Sutter et al. 1999; Zhang and Feng 1998). This term is supported by the frequency extent
of TTI, the number and location of TTI bands, the thresholds of TTI bands, and the lack of lower and/or upper TTI bands in many cells (indicating that spectral integration can remove 2-tone suppression from cortical FRAs). Because the term suppression is no more neutral with regard to mechanism than inhibition (suppression implies basilar membrane 2-tone suppression mechanisms), we, also, use the term inhibition to describe observed reductions of BEF-tone responses with the caveat that this can reflect neural inhibition and/or two-tone suppression.

Responses to the BEF tone were approximated by calculating the mean and standard error of responses from the FRA’s bottom two rows (90 presentations), which essentially consisted of only responses to the BEF tone. Controls with audiovisual monitoring were used to confirm that the low-intensity variable tones did not influence the response. If the estimated BEF-tone activity failed to surpass a criterion value of 0.25 spikes per frequency/intensity combination, the two-tone FRA was reacquired and the responses in the two-tone FRAs were added. This procedure was repeated until the two-tone FRA reached the criterion or we abandoned recording and excluded the FRA from further analysis. For each frequency/intensity combination on FRAs, eight-point weighted smoothing was then performed with nearest neighbors to determine an “average” response (Sutter et al. 1999). Inhibitory regions on the FRA were defined as those where the average response to tone-pairs were <50% of the approximated BEF tone response, i.e., when the variable tone reduced the approximated BEF tone response by >50%.

Data analysis
SPIKE-TIME DISTRIBUTIONS. The timing of individual spikes relative to stimulus onset (spike-times) was measured with millisecond resolution. For each neuron, a histogram of the distribution of spike-times of all responses in the pure-tone FRA was obtained. This can be thought of as a grand-peristimulus time histogram (PSTH), collapsing across all stimuli used in generating the FRA. The shape of each distribution was classified as either unimodal or multimodal. The spike-time distribution was classified as either unimodal or multimodal. The shape of each distribution was classified as either unimodal or multimodal. The “dip” statistic (Hartigan and Hartigan 1985; Hartigan 1985), which reflects the discrepancy between the cumulative frequency distribution and the best-fitting unimodal distribution, was used to determine if a given spike-time distribution deviated significantly from unimodality. P values were determined by comparing the dip statistic of each neuron’s spike-time distribution with the percentage points of the null distribution in Table 1 of Hartigan and Hartigan (1985).

LATENCY VERSUS FREQUENCY. Analyses of the relationship between first-spike latency and frequency were based on the latency maps. Latency maps were similar to FRAs but first-spike latency, rather than response magnitude, was plotted as a function of frequency and intensity. A latency map was constructed for each neuron by taking the median first response latency in a 3 × 3 region at each frequency/intensity coordinate of the FRA. For each point, this region included the two frequency neighbors at the same intensity and three latencies to the same frequencies at the next higher and lower intensities tested. Note that this involves smoothing over several different stimuli and thus differs from the conventional method of computing the mean latency to first spike averaged over a number of responses to the same stimulus. Response latencies <17 ms (unadjusted for 2.65-ms travel time) were prefiltered from the latency maps. Also, if two or fewer frequency/intensity combinations from the 3 × 3 grid elicited responses, that stimulus condition was excluded from the analysis.

The values in the latency map along a fixed intensity provided a latency versus frequency function. For each neuron, several latency/frequency functions were obtained for a range of standard intensities relative to the neuron’s threshold. Threshold and CF were selected from the FRA blindly with respect to the latencies. A population latency/frequency function was obtained by standardizing the axes of the individual latency/frequency functions and computing the median of the functions. The standardization is described in RESULTS. Latency versus intensity functions were also constructed from the latency map by sampling along CF. The intensities of the latency/intensity functions were corrected for the frequency response of the sound-delivery system.

SPECTROTEMPORAL RFs. Spectrotemporal receptive fields (STRFs), were constructed for each neuron. STRFs, representing the average stimulus preceding each spike in the form of a spectrograph, were constructed by a standard “reverse correlation” algorithm that computed the average spectrograph preceding each action potential (DeAngelis et al. 1993; deCharms et al. 1998a). By convention, the STRF time axis is inverted so that the average stimulus is represented in negative time and the time of each spike is defined as zero. Portions of the average stimulus that precede the action potential by a greater amount of time are found at more negative times. The temporal resolution of the STRFs was 1 ms; the frequency resolution was the same as for the pure-tone FRA because these tones were used to derive the STRF. Each STRF was smoothed by a circular Gaussian (standard deviations: 2 ms along the time axis; dependent on the range and number of frequencies used along the frequency axis).

Two types of STRFs, distinguished by how they collapsed data over stimulus intensity, were used. For most recordings, 15 intensities (spaced 5-dB apart) were used. “Intensity-restricted” STRFs were constructed by collapsing around a 20 dB range centered on each intensity. “Grand STRFs” were constructed by collapsing across all intensities.

Histology
To confirm electrode locations, penetrations were marked with two or three electrolytic lesions (~10 μA DC; electrode negative; ~10 s) near the end of five experiments. Animals were transcardially perfused with saline followed by 4% paraformaldehyde. The brains were blocked and stored in 10% sucrose. Fifty-micrometer frozen sections were cut in the horizontal plane, counterstained with cresyl violet, and examined under light microscopy. Electrode penetrations were confined to the caudal bank of the posterior ectosylvian sulcus; only in two animals could we unambiguously identify the laminar position electrode locations. As in earlier studies of PAF (e.g., Heil and Irvine 1998b), recorded neurons were likely distributed over several lamina. In two animals, several tracer injections were made along the dorsal-ventral extent of putative PAF, as defined by the physiological criteria of the present study and Reale and Imig (1980). The resulting pattern of medial geniculate body (MGB) labeling matched the expected distribution of thalamic projections to PAF (Morel and Imig 1987; Rodrigues-Dagaef et al. 1989).

RESULTS
General observations
The general properties of PAF were consistent with those observed in earlier investigations. Compared to A1, PAF neurons responded much less briskly to tones and rapidly habituated to repetition of the same stimulus. Tone-responsive neurons were patchily distributed along the depth of a penetration. PAF was tonotopically organized but had a much coarser tonotopic progression than A1 within the same animal. Intensity-tuned neurons were common and neurons often exhibited late responses in addition to “onset” responses. Increasing the stimulus duration indicated that the late responses were not locked to stimulus offset.

Pure-tone FRAs were obtained for 107 neurons of which 91 had sufficient responses for analysis. Of the 98 neurons for which a two-tone FRA was obtained, only 61 (62%) were
driven reliably enough by the BEF tone to meet the criteria for analysis. Most of the remaining 38% did not respond reliably to two-tone stimuli due to profound habituation to repeated BEF tone presentation despite the use of 2-s interstimulus intervals.

**Pure-tone response area structure**

The majority of neurons (60/91) had an eFRA with only one responsive area, and therefore were classified as single-peaked (Fig. 1, R1). The eFRAs of the remaining 31/91 neurons had at least two response maxima as a function of frequency (Fig. 1, R2). Multipeaked eFRAs had at least two unambiguously distinct excitatory domains (each with clear tips at low intensity) separated by a frequency range that elicited little or no activity (Fig. 1, R2/C1 and C2). These composed 15% (14/91) of the sample. Some neurons had FRAs with diffuse or multiple response maxima as a function of frequency separated by areas of reduced responsiveness (R2/C5 and C6) and were sometimes “beard-like” with a patch of unresponsiveness at higher intensities interrupting an otherwise contiguous region (R2/C3 and C4). The patchy response areas in the latter subtype made firm classification of these neurons difficult as it was not clear whether the patchiness was an actual property of the neurons or an artifact of the low response reliability of the neuron. Therefore we call these “patchy-unclassifiable” FRAs, and they constituted 19% of the sample. Some neurons’ FRAs bordered on classification groups (e.g., the “single-peaked” FRA in R1, C6 shared similarities to patchy-unclassifiable FRAs), suggesting that these properties might be organized along a continuum rather than in discrete classes. Only extremely well-isolated neurons whose multiple response maxima could not be attributable to multiple-unit responses, harmonic distortion, or acoustic resonances were classified as multipeaked or patchy-

![FIG. 1. Pure- and 2-tone frequency/intensity response areas (FRAs) of posterior auditory cortical field (PAF) neurons. In each plot, the x axis is frequency (kHz) and the y axis is intensity (dB). Color indicates the responsiveness at a given frequency/intensity combination, and is normalized to each individual FRA’s maximal response (see following text). The intensity in dB corresponds only to the left-most FRAs, and the intensity scales for the other FRAs have been omitted for esthetics. The intensity scale always spanned 70 dB. Row 1 (R1): pure-tone FRAs of 6 representative single-peaked neurons, which have a single area of responsiveness. R2: pure-tone FRAs of 6 representative multipeaked and patchy-unclassifiable neurons that have \( > 1 \) discernible response area. R3 and R4: each pair of FRAs is the pure-tone and 2-tone FRAs for the same neuron. Darker regions in the 2-tone FRAs are frequency/intensity combinations of the variable tone that reduced the response to the BEF tone (see METHODS for details). Solid white contours in each 2-tone FRA are the outlines of the excitatory regions(s) in that neuron’s pure-tone FRA. The inhibitory band structure for each of the 6 neurons is as follows: R3: LU, MU, and LU; R4: LU, LLUU, and LLMU. The peak firing rate (based on 8-point smoothing; see METHODS) per stimulus presentation in each FRA are as follows (FRAs are based on 1 repetition of each stimulus condition unless otherwise indicated): R1/C1, 2.1; R1/C2, 2.75; R1/C3, 3.05; R1/C4, 3.28 (2 reps); R1/C5, 1.28 (2 reps); R1/C6, 3.11; R2/C1, 2.16 (3 reps); R2/C2, 1.63 (3 reps); R2/C3, 1.40 (2 reps); R2/C4, 4.82 (2 reps); R2/C5, 1.84 (2 reps); R2/C6, 2.32; R3/C1, 2.05; R3/C2, 3.32; R3/C3, 1.4; R3/C4, 2.15; R3/C5, 2.08 (2 reps); R3/C6, 2.50 (2 reps); R4/C1, 1.18 (2 reps); R4/C2, 1.25 (2 reps); R4/C3, 2.1; R4/C4, 2.16; R4/C5, 1.78 (3 reps); and R4/C6, 1.4 (2 reps).]
Spike-time distribution

Generally, a given PAF neuron’s responses to all tested frequency/intensity combinations spanned a large range of times relative to stimulus onset. Representative distributions of spike-times (each spike’s time of occurrence relative to stimulus onset) are shown in Fig. 2. We were interested in whether the spike-time distributions for a given neuron formed a single cluster (i.e., had 1 “mode”) or more than one. A neuron whose dip statistic (Hartigan and Hartigan 1985) was statistically significant ($P \leq 0.05$) was classified as having a multi-modal spike-time distribution (MMSTD; Fig. 2, E–G). Of 91 neurons, 24 met this criterion. Eleven MMSTD neurons had large separations (>5 ms) between spike-time distribution peaks, and the remaining thirteen neurons had jagged peaks separated by <5 ms. The later responses in MMSTD could be due to late spikes following an initial early response or to longer first-spike latency responses. In this respect, it is worth noting that variation in stimulus parameters influences latencies. It is likely that the separation of peaks were blurred (by the variance introduced from using different intensities and frequencies) for many neurons, thereby decreasing the number of reported MMSTD neurons. Counteracting this effect is the possibility that the use of multiple stimuli contributed to MMSTDs. The latter seems unlikely because the relationship between spike-time distribution type (MMSTD or uni-modal STD) and eFRA type (single-peaked, or multi-peaked/patchy-unclassifiable) did not reach significance ($df = 1, \chi^2 = 2.541, P = 0.111$), indicating that there is not a tight relationship between spike-time distribution type and FRA type.

The conservative nature of the dip statistic likely influenced the proportion of neurons classified as having MMSTDs. The conservative nature is due to the null distribution being based on samples from a uniform distribution (Hartigan and Hartigan 1985). This may account for why some neurons that were not classified as having MMSTDs had clearly defined peaks separated by a long time difference (e.g., Fig. 2H). Of the six neurons fitting this profile, most had $P$ values close to 0.05.

Spike-time filtering

In addition to the unusual MMSTDs, PAF neurons also responded over a large range of times. For example the neuron whose spike-time distribution is shown in Fig. 2D could fire action potentials between 30 and 200 ms after tone onset. We wanted to test whether eFRA complexity was preferentially related to later spikes relative to stimulus onset. To assess this, spikes that did not fall within a specified range of time were removed, or “filtered,” from the FRA. The filtered and unfiltered FRAs were then compared. For each pure- and two-tone FRA, the spike-time filter range was gradually narrowed while the effect of each setting was updated graphically. The filter settings used in the examples presented in the following text are optimal in the sense of illustrating described phenomena.

In several multi-peaked neurons, the secondary response areas were clearly due to spikes in the later part of the spike-time distributions. The effect of filtering later responses from eFRAs is illustrated for two neurons in Fig. 3, A and B. In both cases, filtering revealed a simple “early-responsive” area and a more complex “late-responsive” eFRA. The early spikes were confined to a sharply tuned area, while the later spikes were more broadly distributed over the FRA. In 11/14 multi-peaked and 12/17 patchy-unclassifiable neurons, early responses were confined to a narrow range of frequencies in one response peak. Within those subpopulations, 10/11 multi-peaked and 6/12 patchy-unclassifiable neurons had late spikes that were scattered more or less evenly over the entire excitatory re-
Response area. For the remaining neurons in the subpopulation, late spikes were segregated from early spikes and/or were confined to the edges of an individual peak (ignoring threshold effects near each peak’s low intensity eFRA tip). In some MMSTD neurons, there was a correspondence between the peaks in the spike-time distributions and peaks in the FRAs. For example, the FRA in Fig. 3A (left) has the spike-time distribution in Fig. 2H. The spike-time filter cutoff corresponds to the dip between the modes in the spike-time distribution at 40 ms. However, neurons with clear spike-time filtering effects did not necessarily have obvious mode boundaries in the spike-time distribution and vice versa.

The spike-time filtering suggested that secondary response areas of multi-peaked neurons are primarily due to spikes in the later part of the spike-time distribution. Since neurons of all types (including those with single-peaked eFRAs) exhibited a large variation in spike-time distribution, we now ask whether spike-times relate systematically to the parameters of the stimulus that elicited the response.

Response latency versus frequency

Qualitative inspection of the latency maps indicated a clear relationship between latency and tone frequency in many neurons. In these neurons, at supra-threshold intensities the shortest latencies occurred near CF and became progressively longer toward the low- and high-frequency edges of the eFRA, resulting in a V-shaped latency/frequency function (Fig. 4A). To assess this trend quantitatively in the population of neurons, individual latency/frequency functions were standardized as follows. The frequency that corresponded to the minimum latency response (based on the median across all intensities) was designated as the “origin” (arrow in Fig. 4A) and latency/ frequency functions (at several intensities, see following text) were sampled in one-quarter-octave steps, beginning at the origin. Frequency was expressed relative to the origin frequency (in octaves) and latency was expressed relative to the latency at the origin.

For each single-peaked neuron, five latency/frequency functions were obtained from 10 to 50 dB above the neuron’s intensity threshold in 10-dB steps. For each intensity, a population latency/frequency function was derived by taking the median relative latency values of all frequency/intensity combinations from the 60 individual latency frequency functions (see Fig. 5A legend for more details). The relative latency increased with shifting frequency from the origin. Collapsing over intensity, the median of the five population functions had a V shape (Fig. 5B). For the population of single-peaked neurons in PAF (Fig. 5B), latency shifts with frequency (Table 1) were statistically significant (linear regression analysis: lower limb: \( r = -0.995, F = 302.67, P = 0.0004, \text{slope} = -14.28 \text{ms/octave}; \) upper limb: \( r = 0.991, F = 275.65, P < 0.0001, \text{slope} = 8.3 \text{ms/octave} \)).

The individual/latency frequency functions of the multi-peaked and patchy-unclassifiable neurons tended to be discontinuous due to the areas of unresponsiveness between response peaks. However, these discontinuities appeared to be averaged in an analysis of this subpopulation of 31 neurons. For these neurons (Fig. 5, C and D), latency shifts with frequency were statistically significant (linear regression: lower limb: \( r = -0.963, F = 101.58, P < 0.0001, \text{slope} = -12.28 \text{ms/octave}; \) upper limb: \( r = 0.950, F = 83.43, P < 0.0001, \text{slope} = 11.59 \text{ms/octave} \)). A separate analysis of the 14 multi-peaked neurons was also statistically significant (lower limb: \( r = -0.961, F = 83.5, P < 0.0001, \text{slope} = -17.14 \text{ms/octave}; \) upper limb: \( r = 0.985, F = 195.95, P < 0.0001, \text{slope} = 11.65 \text{ms/octave} \)). As were the results for the 17 patchy-unclassifiable neurons (lower limb: \( r = -0.974, F = 145.31, P < 0.0001, \text{slope} = -11.25 \text{ms/octave}; \) upper limb: \( r = 0.895, F = 36.4, P = 0.0002, \text{slope} = 8.45 \text{ms/octave} \)).

The standardization of individual latency/frequency functions introduced an inherent bias, necessitating a guarded interpretation of these correlations. The standardization guarantees that at least one of the five individual latency/frequency functions for a given neuron will have the origin at the function minimum. This could conceivably bias the high- and low-frequency limbs of the latency/frequency function to slope upward or downward, respectively. To determine if this bias, by itself, could account for statistically significant correlations...
between latency and frequency, the following Monte Carlo analysis was done using the single-peaked neuron data. For each individual neuron’s latency/frequency function, the relative latencies on each side of the origin were randomly shuffled while leaving the value at the origin unchanged. One thousand shuffled versions of each neuron’s five latency/frequency func-

FIG. 4. Analysis of 1st-spike latency in a single- and multipeaked neuron. A, left: latency map of a single-peaked neuron with latency rendered in pseudocolor. The correspondence between the color and latency (in ms) is indicated in the color scale to the right of the plot. The arrow along the x axis indicates the frequency origin (see text for details). Points along the dashed horizontal line in the latency map were sampled in quarter-octave steps and used to derive the neuron’s latency/frequency function. A, right: the latency/frequency function for the same neuron, at the intensity indicated by the dashed line in the latency map. Frequencies along the x axis are expressed relative to the frequency origin (in octaves). The y axis on the left side of the plot is the latency. The y axis on the right side is the latency relative to the latency at the origin. [In this neuron, the frequency origin and character frequency (CF) are the same, but this correspondence was not observed in all neurons.] B, left: latency map of a multipeaked neuron. The 2 dashed vertical lines indicate the locations of the 2 lowest threshold response regions, denoted CF_L and CF_H. The latency map values along these lines are plotted in the accompanying latency/intensity functions (B, right).

FIG. 5. A and B: relative latency vs. relative frequency in the population of 60 single-peaked neurons. Conventions are as in the line plot in Fig. 4A, right. A: each line plot is a population latency/frequency function derived from the 60 individual latency/frequency functions for a given intensity relative to each individual neuron’s threshold. Each data point is the median of the values of the 60 individual latency/frequency functions at the corresponding relative frequency. Because each neuron’s individual latency/frequency function could be defined over different relative frequency ranges (because each neuron had a different breadth of frequency tuning), the number of observations underlying each data point is variable. If <8 observations were available at a given relative frequency/intensity combination, the data point was omitted from the analysis (e.g., 1.5 octaves at 10 dB). The functions for 20 and 40 dB, omitted for clarity, exhibited similar trends. B: the median of the 5 population latency/frequency functions at 10, 20, 30, 40, and 50 dB. Lines demarcate regression fits for each limb of the function (statistics in text). The regression model contained no y intercept and excluded the data point at the origin. C and D: relative latency as a function of relative frequency in a population of 31 multipeaked or patchy-unclassifiable neurons. Graphs are analogous to A and B.
tions were generated. From these functions, 1,000 population latency/frequency functions were generated by the same procedure used to derive the population functions in Fig. 5B. A correlation coefficient was computed for each limb of each simulated population function. This yielded two reference distributions, each comprising 1,000 correlation coefficients.

Comparisons of the actual correlations against the reference distributions strongly suggest that the correlations are not simply a standardization procedure artifact. For the high-frequency limb of the population latency/frequency function, the correlation coefficient of 0.974 exceeded all of the coefficients in the corresponding reference distribution (i.e., $P < 0.001$). For the lower limb, the correlation coefficient of $0.987$ was stronger than 98.7% of the reference distributions, corresponding to a $P$ value of 0.013.

**Inter-CF latency comparison in multipeaked and patchy neurons**

We now consider whether there were consistent differences in the latencies for the multiple bands within multipeaked and patchy-unclassifiable neurons. We obtained latency versus intensity functions using the procedure illustrated in Fig. 4B. A multipeaked neuron’s latency map is shown in the pseudocolor plot, and the latency versus intensity functions at the higher and lower characteristic frequencies ($\text{CF}_H$ and $\text{CF}_L$, respectively) are shown in the accompanying line plot. $\text{CF}_L$ and $\text{CF}_H$ were identified blindly with respect to response latency. If a neuron had more than two response peaks/patches, $\text{CF}_L$ and $\text{CF}_H$ were set to the CFs of the two lowest threshold response peaks. Latency/intensity functions were obtained for the 14 multipeaked and 17 patchy-unclassifiable neurons. In general, latency/intensity functions were diverse, both in the terms of shapes of individual functions and in the relationship between the functions at $\text{CF}_H$ and $\text{CF}_L$ (Fig. 6).

There was no obvious correspondence in the shape or slope of the two latency/intensity functions for a given neuron. In some cases, a U-shaped function at one CF was accompanied by a monotonically decreasing function at the other CF (Fig. 6F), and functions sometimes crossed at higher intensities (Fig. 6D and J). Often, the latency at one CF was clearly shorter than the latency at the other CF. The magnitude of this latency difference varied as a function of intensity, but the direction of the difference was usually maintained (Fig. 6A, B, E, F, and G–I).

In some neurons, it was clear that lower threshold response peak latencies were shorter than higher threshold response peak latencies (e.g., Fig. 6B, D, and E). This observation is consistent with the spike-time filtering results (Fig. 3) in which removing later spike-times tended to eliminate higher threshold response peaks. To address this quantitatively, for each neuron, two latency measures were obtained by taking values from its $\text{CF}_H$ and $\text{CF}_L$ latency/intensity functions at the same intensity relative to the threshold of each individual response peak. So, for example, the $L_{10\text{rel}}$ measure for the neuron with intensity-latency functions from Fig. 6B was taken at 12 dB for the $\text{CF}_L$ peak and 32 dB for the $\text{CF}_H$ peak. A paired $t$-test was then used to compare the high- and low-threshold latency measures of all

| TABLE 1. Slope magnitude of frequency/time relationship |
|---------------------------------|--------------------|-----------------|--------------------|-----------------|
| Population latency vs. frequency† (single peaked)  | NA                 | NA              | 70                 | 120             |
| Population latency vs. frequency (multipeaked and patchy-unclassifiable)  | NA                 | NA              | 81                 | 86              |
| Median STRF slope‡          | 173                | 156             | 56                 | 63              |

NA, not applicable. * Upward corresponds to up-sloping spectrotemporal receptive fields (STRFs) and the lower limb of the population latency versus frequency function. † Note that kHz/s and octaves/s are frequency change per unit time and, therefore, are the inverses of the slopes of the Fig. 5 plots, where latency is on the ordinate and frequency is on the abscissa. ‡ For STRFs, the median slope is only calculated for cells classified as sloping (50.5%). If the STRF was V shaped, both limbs were used. Slopes were derived by visually fitting a line to the STRFs.

**FIG. 6.** Latency vs. intensity relationships of different response peaks in 10 multipeaked or patchy-unclassifiable neurons. Each plot depicts the latency as a function of intensity at the CF of the 2 lowest threshold response areas in each neuron. ●, data for the low-frequency peak ($\text{CF}_L$); ×, $\text{CF}_H$ data. Thresholds at each CF are indicated by the x axis coordinate of the leftmost data point of each function.
the multipeaked and patchy-unclassifyable neurons (excluding 4 neurons that had identical thresholds at CF_H and CF_L). A separate test was carried out for each latency measure from threshold (L_0rel) to 50 dB above threshold (L_50rel) in 5-dB steps. The distribution of differences of the L_10rel measure (high-threshold L_10rel – low-threshold L_10rel) is shown in Fig. 7A (mean difference = 10.47 ms, t = 2.77, df = 26, P = 0.0101). The 25th, 50th, and 75th percentiles of the distribution were 3.3, 12, and 20.6 ms, respectively. Statistically significant positive differences were also found for the L_0rel, L_5rel, and L_15rel measures, with mean differences between 8.2 and 10.5 ms. For the remaining latency measures (L_20rel through L_50rel), the mean latency difference was consistently positive (between 3.6 and 13.6 ms), but the results were not statistically significant. This lack of statistical significance may be due to the limited intensity range of higher threshold latency/intensity functions which resulted in a fewer number of observations in these analyses.

Similar analyses were also performed separately for multi-peaked and patchy-unclassifyable neurons. For multipeaked neurons, a paired t-test of the high- and low-threshold L_10rel was statistically significant (mean difference = 10.65 ms, t = 2.81, df = 12, P = 0.0158). Statistically significant results in the same direction were also obtained for L_5rel, L_15rel, and L_30rel. For the remaining latency measures, the mean high-threshold – low-threshold latency difference was consistently positive. For the patchy neurons, none of the analyses reached statistical significance; however, the direction and magnitude of the mean latency difference was consistent with the overall pattern of results for multipeaked neurons.

Another question is whether the latency for the low-frequency peak (in contrast to the low-threshold peak) is consistently shorter (or longer) than at CF_H. To address this quantitatively, the latency difference was computed with respect to frequency peaks (e.g., high-frequency L_10rel – low-frequency L_10rel). The median L_30rel difference was 7.1 ms and L_10rel difference was 8.3 ms (Fig. 7B). The median and mean of the distributions of the latency differences were positive for all latency measures between L_0rel and L_35rel inclusively, suggesting that the CF_H had longer latencies than CF_L. However paired t-tests did not reach significance (e.g., for L_10rel mean difference = 3.95 ms, t = 0.996, df = 30, P = 0.327). There was a wide range of frequency separation between CF_H and CF_L. The distribution of the CF_H/CF_L ratio, shown in Fig. 7C, had a median of 2.42. For multipeaked neurons, the median CF ratio was 2.32, exceeding the median CF ratio of 1.56 observed for dorsal A1 multipeaked neurons (Sutter and Schreiner 1991). The median CF ratio in patchy neurons was 2.78.

Spectrotemporal RFs

Naturally occurring sounds vary in the temporal as well as the spectral domain. In this context, it is important to characterize PAF neurons’ STRFs. STRFs also complement the previous first-spike latency and spike-time analyses by systematically relating the times of all spikes to frequency. Pure-tone STRFs were obtained for 91 neurons. Neurons were classified based on grand STRFs, which collapsed the data for responses over all tested intensities (e.g., Fig. 8, A–H). STRFs were classified by visual inspection, blindly with respect to the neuron’s FRA. To avoid graphical display biases, initial assessment of all STRFs was performed with a 50-ms time axis and 4-octave frequency scale.

Despite the marked heterogeneity of STRFs, all neurons could be classified into one of five types based on the configuration of maximal response areas (light blue to red, excluding dark blue in Fig. 8). “Simple” STRFs (Fig. 8, A and B) had one or possibly two closely spaced blob-shaped response maxima. In a significant proportion of simple STRFs, the blobs were somewhat elongated along the vertical or horizontal dimension (Fig. 8B). A second STRF class had positive sloping (“up-sloping”) response maxima (Fig. 8, C and D). Within this class, there was variation with respect to a few notable features. One feature was the slope magnitude (Table 1); another was whether the response maxima consisted of a single elongated “ramp” (Fig. 8C) or a “staircase” with two or more distinct steps (Fig. 8D). Ramps were also sometimes uneven or discontinuous, giving them a “patchy” appearance. A third class had negative sloping maxima (“down-sloping”) and is directly analogous to up-sloping STRFs. Down-sloping STRFs also had both ramp (Fig. 8E) and staircase (Fig. 8F) features.

A fourth STRF class exhibited both up- and down-sloping
features that combined, forming a V-shaped structure (Fig. 8, G and H). In most cases, one limb of the V was clearly dominant. V-shaped STRFs could have any combination of ramp, patchy ramp, or staircase features. The up-sloping, down-sloping, and V-shaped categories can all generally be thought of as “sloped” STRFs. The occurrence of ramp or staircase features in sloping STRFs related straightforwardly to FRAs with ramp features associated with contiguous response areas and staircases associated with multi-peaked and disjointed response areas. The fifth STRF type had a complex appearance with three or more disjointed maxima, but otherwise had no common structure. Some complex STRFs appeared to have an incipient sloped structure that was not convincing enough to justify classification in any category. About half (50.5%) of the grand STRFs were sloping with the down-sloping, up-sloping, and V subtypes constituting 24, 39, and 37%, respectively, of the sloping STRFs. Simple and complex STRFs made up 35.2 and 14.3% of the sample, respectively.

The relationship between spectrotemporal properties and stimulus intensity was probed by constructing STRFs that were confined to a narrow range of stimulus intensities (in contrast to the grand STRFs, which incorporate responses to all tested intensities). This yielded several “intensity-restricted” STRFs for each neuron, with each one representing a 20 dB range centered at a different intensity. In the lower part of Fig. 8, each row of panels shows four STRFs for a single neuron: the grand STRF followed by three intensity-restricted STRFs at low, mid, and high intensities. The intensity-restricted STRFs were classified in the same manner as the grand STRFs. Although the percentage of neurons with at least one sloping intensity-restricted STRF (53/91, 58%) was similar to the percentage of sloped grand STRFs, there were some differences between grand and intensity-restricted STRFs. In general, the structure of individual intensity-restricted STRFs were less ambiguous than the grand STRFs. For example, the grand STRF of Fig. 8I was a marginal example of a down-sloping staircase. But the mid-intensity STRF (Fig. 8K) had a much clearer staircase structure that is obscured when collapsing

![Image of Fig. 8](http://example.com/fig8)
across intensities. If sloped structure was present, it was usually evident at several intensities, (note that the intensity-restricted STRFs presented here are selected from a larger series, and there are therefore “gaps” in the presented sequence). In general, ramp features were more elongated in the intensity-restricted STRFs than in the grand STRFs (e.g., low and mid intensities in Fig. 8, N and O) and staircases were more distinct (e.g., high intensities in Fig. 8). Nearly all the neurons with V-shaped grand STRFs showed alternating dominance of the upward and downward limbs over a sequence of intensities.

Two-tone response area structure

The two-tone FRAs of PAF neurons exhibited diversity in the number, size, and shape of inhibitory bands. Examples of two-tone FRAs with various inhibitory band structures are presented in Fig. 1, R3 and R4. To allow for comparisons with two-tone results in A1, inhibitory band structure classification followed the methods of Sutter et al. (1999). Briefly, each inhibitory band was assigned one of three labels depending on its location with respect to the excitatory frequency range. Inhibitory bands extending below or above the excitatory frequency range were labeled as a lower band (L) or upper band (U), respectively. A band whose frequency extent was contained entirely within the pure-tone eFTC was labeled a “middle” (M) band. Note that an inhibitory band lying between two excitatory peaks would be labeled a middle band (e.g., Fig. 1, R3/C3 and C4). The entire inhibitory band structure was classified by concatenating individual bands labels so that a neuron with the one lower and one upper inhibitory band would be classified as LU. A neuron with two distinct lower and two middle inhibitory bands would be classified as LLMM. Neurons within the same class often varied greatly in the properties of individual bands. In cases where the inhibitory structure was uncertain, the rating defaulted to the simplest possible classification (where the structure closest to LU would be considered simpler).

In both A1 and PAF, the largest single class of inhibitory band structure was LU, which is how one might have historically conceptualized inhibitory bands (Fig. 9). The incidence of the LU structure was significantly lower in PAF than in A1 (19 vs. 38%, $\chi^2 = 6.651$, df = 1, $P < 0.01$). Comparisons with dorsal and ventral A1 (A1v and A1d) revealed that this was due to differences between PAF and A1v, where 50% of the neurons had LU structure, while only 16% of A1d neurons had LU structure (Fig. 9). In both A1 and PAF, the incidence of LU inhibitory band structure is likely an overestimate due to the convention of adopting simpler classifications in cases of ambiguity. For non-LU band structure, PAF exhibited a larger range in complex band structure type than A1. In general, PAF’s inhibitory band structure is more complex than A1v’s and at least as complex as A1d’s.

Generally, strongly intensity tuned neurons with circumscribed eFRAs had broad flanking inhibitory bands that merged at CF (e.g., Fig. 1, R3/C1 and C2). In most neurons, inhibitory bands flanking excitatory areas had thresholds closely matching the excitatory area; however, secondary inhibitory bands
further away from the pure-tone response areas were generally higher threshold and less clearly defined.

Relationship of spike time and iFRA complexity

Inhibitory band structure of spike-time filtered two-tone FRAs did not necessarily match those of un-filtered ones. Twenty-three percent (14/61) of neurons’ inhibitory band structure classification was altered by spike-time filtering. Changes in inhibitory band classification were sometimes precipitated by pure-tone FRA spike-time filtering. For instance, a middle inhibitory area interposed between response peaks of a multipeaked neuron would be redesignated as either a lower or upper band when the long latency, off-CF pure-tone response area was eliminated. This occurred in five neurons, and is illustrated for two neurons in Fig. 10, A and B. In four neurons, spike-time filtering transformed the classification to LU, increasing the proportion of neurons with LU inhibitory band structure to 31%. Elimination of late responses from the two-tone FRA also could change inhibitory band structure. In some cases, as in Fig. 10C, bands became much more apparent by filtering late responses that were scattered in the interior of the bands. It is unlikely that these late responses were spontaneous discharges because spontaneous activity was extremely low. Some inhibitory bands broadened when long latency responses on the borders were removed. This sometimes caused two narrowbands to merge into one larger band. In two neurons, one of which is illustrated in Fig. 10D, a new inhibitory band appeared flanking the excitatory area. In general, these latency-related phenomena indicated a complex interaction between spectral and temporal coding.

DISCUSSION

Comparison with earlier studies of PAF

This investigation augments the understanding of PAF in three major respects. First, it further demonstrates complexity in PAF eFRAs beyond narrow and V-shaped frequency tuning. Second, it reveals that spectral and temporal response properties are intricately interrelated and the potential importance of the timing of PAF responses. Third, it provides the first direct

FIG. 10. Raw, unsmoothed FRAs demonstrating the effect of filtering the long latency responses in 6 neurons (A–F). The length of the line at each FRA coordinate is proportional to the number of spikes elicited by that frequency/intensity combination. Left: unfiltered version of FRA. Middle: the same FRA in left with later spikes removed. Right: the same FRA in left with earlier spikes removed. The temporal window of the spike-time filter is indicated in the bottom right of the plots in middle and right. A and B: pure-tone FRAs of 2 neurons with superimposed outlines of the inhibitory bands. C and D: 2-tone FRAs of 2 neurons.
evidence of inhibition shaping PAF eFRAs. In this section, we will compare the present results to earlier studies with particular attention to methodology that may contribute to differences between reported results. Table 2 compares methodological parameters of several pertinent PAF investigations, including studies of frequency tuning using tones (Heil and Irvine 1998b; Phillips and Orman 1984) and of FM using sweeps (Heil and Irvine 1998a; Tian and Rauschecker 1998). For comparison purposes, we have excluded studies primarily focusing on CF response properties.

**SPECTRAL PROPERTIES.** A major finding of this study is the large diversity of eFRAs types in PAF and that 33% of PAF neurons did not have simple, single-peaked eFRAs. This is in general agreement with Heil and Irvine (1998b), who reported similar response area diversity with 25% of PAF neurons possessing broad or complex shapes. In contrast, Phillips and Orman (1984) described PAF neurons’ eFTCs as narrow and V shaped. Interestingly, the latter study only included neurons responding with latencies <50 ms, while studies reporting complex eFRAs (i.e., the present study; Heil and Irvine 1998b) did not exclude neurons on the basis of latency. This is consistent with Reale and Imig’s (1980) observation that early and late responses of neuron clusters in auditory cortex were associated with narrow and broad frequency tuning, respectively. Some of the present study’s results, such as the latency versus frequency relationship and spike-time filtering effects, extend this concept further by revealing that the relationship between spike-time and spectral complexity is inherent in the single neuron receptive field organization.

Several factors could have led to the relatively high estimate of neurons with multipeaked and patchy-unclassifiable eFRAs in this study (Table 2) (note the same factors could relate to the high estimates of complex inhibitory band structure). For example, because we did not calibrate at the tympanic membrane, filtering by the pinna could contribute to the observance of multipeaked neurons. This probably was not a large contributor as, within the same animal, frequencies that were located between peaks in one cell could be the most sensitive in another.

The trade-off between number of different stimuli presented and total repetitions of each stimulus also likely contributed to the high percentage of multipeaked and patchy-unclassifiable eFRAs. Compared to other studies (e.g., Heil and Irvine 1998b), the present study sampled frequency more finely over a broader range at the expense of fewer stimulus repetitions. Finer sampling over a wider range could reveal response areas that would have otherwise been missed with coarser sampling as suggested by several very sharply tuned areas within multipeaked eFRAs (e.g., Fig. 1, R4/C5). We also encountered a few multipeaked neurons that were tested with a frequency range of five octaves where some peaks would go undetected with a narrower stimulus frequency range (e.g., see Fig. 1, R3/C3). Therefore by using more frequencies we are more likely to correctly detect multiple response areas.

While the two examples demonstrate how the advantages of using more frequency/intensity combinations could lead to higher estimates of multipeaked and patchy-unclassifiable eFRAs, using fewer repetitions introduces a disadvantageous effect, which might inflate the reported percentages of patchy-unclassifiable neurons. By presenting fewer stimulus repetitions, the false identification of a patchy eFRA becomes more likely because of unreliable responsiveness to any single stim-

### TABLE 2. Comparison of methodological parameters in five investigations of PAF

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recording location</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lamina</td>
<td>Middle and lower layers</td>
<td>All layers w/possible bias toward middle</td>
<td>All layers w/possible bias toward middle</td>
<td>All layers w/possible bias toward middle</td>
<td></td>
</tr>
<tr>
<td><strong>Maximum penetration depth, µm</strong></td>
<td></td>
<td>6000</td>
<td>3300</td>
<td></td>
<td>6000</td>
</tr>
<tr>
<td><strong>Stimulus parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery</td>
<td>Sealed/binaurally optimal</td>
<td>Sealed/binaurally optimal</td>
<td>Free field—90° azimuth</td>
<td>Free field—0° azimuth</td>
<td>Sealed/binaurally optimal</td>
</tr>
<tr>
<td>Type</td>
<td>Pure tone</td>
<td>Pure tone</td>
<td>Pure tone</td>
<td>Pure tone</td>
<td>FM/linear sweep velocity</td>
</tr>
<tr>
<td><strong>Duration, ms</strong></td>
<td>100</td>
<td>200</td>
<td>50</td>
<td>20</td>
<td>Varied with sweep velocity (&lt;3 s)</td>
</tr>
<tr>
<td><strong>Repetitions</strong></td>
<td>High (30†)</td>
<td>High (5–15)</td>
<td>Low (1–3)</td>
<td>20</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Frequency sampling density/scale</strong></td>
<td>NR/log scale</td>
<td>Low (8–20)/linear scale</td>
<td>High (45–80)/log scale</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Shaping</strong></td>
<td>5 ms linear</td>
<td>8.5 ms cosine-squared</td>
<td>3 ms linear</td>
<td>5 ms linear</td>
<td>5 ms linear</td>
</tr>
<tr>
<td><strong>Presentation rate/order</strong></td>
<td>0.7–3 s/NR</td>
<td>1 s/pseudorandom</td>
<td>0.4–2 s/pseudorandom</td>
<td>≥1 s/NR</td>
<td>NR/pseudorandom</td>
</tr>
<tr>
<td><strong>Latency criterion,</strong> ms</td>
<td>≤50</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>Gas</td>
</tr>
<tr>
<td><strong>Anesthetic</strong></td>
<td>Pentobarbital sodium</td>
<td>Pentobarbital sodium or ketamine</td>
<td>Pentobarbital sodium</td>
<td>Pentobarbital sodium or ketamine</td>
<td>Complex sounds</td>
</tr>
<tr>
<td><strong>Search stimuli</strong></td>
<td>Pure tones</td>
<td>Pure tones, clicks, and noise</td>
<td>Pure tones</td>
<td>Pure tones, clicks, and noise</td>
<td></td>
</tr>
</tbody>
</table>

NA, not applicable; NR, not reported explicitly. PAF, posterior auditory field; eFTC, excitatory frequency tuning curve. † Inferred from Fig. 2 in Tian and Rauschecker (1998). Also, in this study the penetrations were oriented dorsoventrally. ‡ This estimate is based on Fig. 1 in Phillips and Orman (1984). § This refers to the response latencies of the neurons that were included in the study.
SPECTROTEMPORAL PROPERTIES. PAF neurons have been hypothesized to encode signals varying slowly over time (Heil and Irvine 1998a,b; Tian and Rauschecker 1999) and the rate of FM sweeps (Tian and Rauschecker 1998). It should be noted that while forms of FM other than ramps (e.g., pulsed FM, transient FM embedded in tones, and periodic FM) are important, because studies of PAF have been limited to FM ramps, our discussion will focus on this form of FM. The present analysis of sloping STRFs could relate to FM rate sensitivity. However, we did not test our cells with actual FM sounds and the predictive power of STRFs—derived in different ways—to responses to arbitrarily complex sounds is controversial (e.g., see deCharms et al. 1998b; Keller and Takahashi 2000; Poon and Yu 2000; Theunissen et al. 2000). For example, nonlinear summation evoked by FM cannot always be predicted by responses to pure-tone stimuli (e.g., Erulkar et al. 1968; Fuzessery 1994; Suga 1968; Whitfield and Evans 1965), and only the onset of our stimuli may evoke envelope effects (Heil and Irvine 1998b; Phillips et al. 1996) differing from those in the FM studies. Therefore the following section should be considered speculative.

Despite these potential pitfalls, our results are roughly consistent with the notion that the slopes of STRFs relate to FM rate preference. Tian and Rauschecker (1998) found that ~70% of neurons preferred FM rates <200 kHz/s, while Heil and Irvine (1998a) found that the majority of PAF neurons preferred rates <100 kHz/s. (Some of the methodological differences between these FM studies are shown in Table 2.) The median STRF slope in the present study of 170 kHz/s is in the range of the other studies’ FM rate preference. The steeper STRF slopes of the present study, in comparison to Heil and Irvine (1998a), may be explained by two considerations: first, neurons with higher CFs have faster FM rate preferences, as consistent with the notion that the slopes of STRFs relate to FM rate sensitivity. However, we did not test our cells with actual FM sounds and the predictive power of STRFs—derived in different ways—to responses to arbitrarily complex sounds is controversial (e.g., see deCharms et al. 1998b; Keller and Takahashi 2000; Poon and Yu 2000; Theunissen et al. 2000). For example, nonlinear summation evoked by FM cannot always be predicted by responses to pure-tone stimuli (e.g., Erulkar et al. 1968; Fuzessery 1994; Suga 1968; Whitfield and Evans 1965), and the onset of our stimuli may evoke envelope effects (Heil and Irvine 1998b; Phillips et al. 1996) differing from those in the FM studies. Therefore the following section should be considered speculative.

Despite these potential pitfalls, our results are roughly consistent with the notion that the slopes of STRFs relate to FM rate preference. Tian and Rauschecker (1998) found that ~70% of neurons preferred FM rates <200 kHz/s, while Heil and Irvine (1998a) found that the majority of PAF neurons preferred rates <100 kHz/s. (Some of the methodological differences between these FM studies are shown in Table 2.) The median STRF slope in the present study of 170 kHz/s is in the range of the other studies’ FM rate preference. The steeper STRF slopes of the present study, in comparison to Heil and Irvine (1998a), may be explained by two considerations: first, neurons with higher CFs have faster FM rate preferences, as reported in units of kHz/s (Heil and Irvine 1998a; Tian and Rauschecker 1998). And second, the median CF of neurons with sloped STRFs was 6.8 kHz while the median CF in the Heil and Irvine study was half that. Consistent with the possible relationship between STRF slope and FM rate preference, the CF and STRF slope magnitude (in kHz/s) of the neurons in this study were positively correlated (Spearman rank test, CF vs. upward slope: \( r_s = 0.333, P < 0.05 \); CF vs. downward slope: \( r_s = 0.693, P < 0.0001 \)). However, this correlation must be interpreted cautiously because the stimulus frequency axis in this study was sampled logarithmically not linearly. This is pertinent because the progressively coarser sampling in kHz at higher frequencies would make it more difficult to resolve shallow frequencies (in units of kHz/s) at higher frequencies.

In addition to the analysis of individual neurons’ STRFs, the present study used a population analysis (e.g., Fig. 5) to address the question of frequency/latency slope in octaves/s. The STRF and population analyses yielded similar median upward slope magnitudes, but the population frequency/latency analysis of single peaked neurons had nearly double the downward slope of the STRF analysis (Table 1). However, this discrepancy is not surprising given that the STRF analysis was confined to neurons with sloped STRFs. In contrast, the population analyses also included neurons with simple, vertically oriented STRFs whose response latency (time on the STRF) is more or less constant as a function of frequency. The slope (\( \Delta \text{frequency}/\Delta \text{time} \)) of such a vertically oriented STRF would approach infinity, and therefore inclusion of these neurons in the population analysis would lead to steeper slope estimates. Based on the overall pattern of results, we hypothesize that temporal shifts in excitation as a function of stimulus frequency may contribute to the FM rate tuning of PAF neurons. This hypothesis can be tested in the future by combining the techniques of the present study with presentations of temporally varying stimuli and comparing the two types of data in the same neurons. Since inhibitory components likely also contribute to FM rate and/or direction selectivity (Fuzessery 1994; Fuzessery and Hall 1996; Shamma et al. 1993; Suga 1965a,b), excitatory and inhibitory domains may shape FM rate preference in a complementary manner.

Temporal complexity of PAF responses

PAF neurons fired action potentials to pure tones over a broad range of times. Earlier studies referred to late responses in PAF cluster activity but didn’t examine late responses in detail (Reale and Imig 1980). The present study shows that both early and late responses are produced by single, well-isolated neurons. Additionally, late responses may contribute to MMSTDs in PAF neurons. It is possible that some late responses were offset responses, but we observed that late responses were usually not locked to stimulus offset in accordance with earlier studies (e.g., Phillips and Orman 1984). Therefore the late responses in PAF seem to clearly distinguish it from A1.

It is interesting to speculate on potential functional roles of late responses. These responses could play a role in encoding sounds with time-varying frequency components such as FM. They also could play a role in combining spectral information across features of segmented sounds, such as vowel-consonant transitions. Late responses might also contribute to duration tuning by facilitating offset responses for durations where the two response components overlap in time (Brand et al. 2000; Casseday et al. 1994; He et al. 1997; Narins and Capranica 1980). Another interesting possibility is that there exists a time-multiplexed code with early responses finely encoding frequency, while later responses encode something else. Within this context it is intriguing that bat single auditory cortical neurons have been found that have distinct response properties for echolocation and passive listening behaviors (Ohlemiller et al. 1996; Razak et al. 1999). This suggests that convergence of information from parallel pathways and mul-
tplexed behaviorally relevant codes might be important auditory cortical properties.

The present study has confirmed the presence of U-shaped latency versus intensity functions in PAF neurons. Previously, Phillips et al. (1995) interpreted U-shaped latency/intensity functions as a possible artifact of spontaneous activity (see also Phillips and Orman 1984), but Heil and Irvine (1998b) have since also reported U-shaped functions in PAF. This contrasts reports of solely monotonically decreasing latency/intensity functions in A1 (e.g., Calford and Semple 1995). U-shaped functions may reflect sensitivity to both rising and falling transients, a property that might be advantageous for processing amplitude modulated signals (Schreiner and Urbas 1988).

Potential mechanisms underlying spectral and temporal characteristics of PAF neurons

The present study’s experimental paradigm is limited in addressing mechanism(s) of STRF formation. This difficulty is rooted in the overlapping gradients of inhibitory and excitatory inputs throughout much of the frequency/intensity response area (Caspar et al. 1994; Greenwood and Maruyama 1965; Jen and Zhang 2000; Palombi and Caspary 1996; Sutter et al. 1999). Stated differently, a pure tone evokes multiple, independent excitatory and inhibitory afferent signals, shaped at all levels of the ascending auditory system (Phillips and Hall 1992). The inputs combine to produce a net effect for each frequency/intensity combination that likely varies along a continuum from strong inhibition to strong excitation. We measured only this net effect; therefore discussion of the underlying mechanism must be considered speculative. Nevertheless, such discussion is crucial for formulating testable hypotheses concerning the organization of the system.

Why are there fewer neurons with classical LU (surround) inhibitory structure in PAF than A1? PAF neurons might have several distinct inhibitory regions separated in frequency. It is also possible that several distinct excitatory or two-tone facilitatory regions separated in frequency create middle, multiple upper, and/or multiple lower inhibitory bands by “splitting” inhibitory bands (Sutter et al. 1999). We cannot distinguish between these two possibilities with the techniques employed in this study. But with either mechanism, at some stage(s) in the ascending auditory system, excitatory and/or inhibitory inputs with large frequency separation might be integrated to produce the complex inhibitory band structure observed in the present study. In view of the similar percentages of LU neurons in PAF and dorsal A1 (A1d) (Sutter et al. 1999), and the interconnections among auditory cortical fields (Bowman and Olson 1988; Imig and Reale 1980; Rouiller et al. 1991), it is tempting to speculate that PAF complexity is conferred by inputs originating in A1d. However, this is unlikely. First, ablation of ipsilateral A1 indicates that serial connectivity between A1 and PAF is not required to explain some basic properties (intensity tuning, latency, threshold, or discharge rate) of short-latency PAF responses to CF tonal stimuli (Kitzes and Hollrigel 1996). It should be noted, however, that the complexity we report is mainly due to integration beyond CF. Second, the separation of excitatory peaks in PAF is wider than in A1d [median CF ratio in PAF: ~2.4, present study; in A1d: ~1.5, Sutter and Schreiner (1991)]. Therefore PAF and A1d likely derive complex inhibitory band structure in parallel with PAF complexity derived from integrating multiple excitatory and/or inhibitory inputs with large frequency separation.

Another question is what causes some PAF neurons’ excitatory and inhibitory areas to appear patchy and indistinct in comparison to A1 neurons? One possibility is that a greater extent of overlap of inhibition and excitation leads to overall weak net excitation and potentially less reliable responses, in some parts of the eFRA. Increased jitter in the relative timing of excitation and inhibition could also contribute to the unreliability of responses in the following manner. If the excitatory input arrives before the inhibitory input, a response could be generated before the inhibition takes effect; on the other hand, inhibition arriving prior to excitation could preempt a response. Thus jitter in the relative timing of inhibition and excitation would introduce variability in whether the neuron responds to a given stimulus. A similar argument could be made with respect to jitter in the duration of inhibition since inhibition of longer duration would more effectively prevent a late response.

A related question is what are the sources of input and potential mechanisms that underlie the late responses in PAF? Two pertinent neural mechanisms are: 1) distinct anatomical pathways with differing delays (Carr 1993; Carr and Konishi 1990; Kuwabara and Suga 1993) and 2) intrinsic cellular or synaptic properties that may delay or prolong neural responses through inhibition or slow kinetics (Saïtoh and Suga 1995; Zhang and Oertel 1994). These mechanisms are not mutually exclusive, as different biophysical properties often segregate along distinct anatomical pathways (Metherate and Cruikshank 1999; Zhang and Oertel 1994). Regarding the second mechanism, direct tests of the intrinsic properties of PAF neurons (or known afferent neurons to PAF) have not been made; however, the prevalence of inhibition in PAF suggests that inhibition in the afferent pathway to (or within) PAF delays responses. Regarding the first mechanism, later responses may arise from descending influences from classical auditory or association areas (Irvine and Huebner 1979), the limbic system (Prieto and Winer 1993), or recurrent feedforward or feedback circuitry, all of which may entail polysynaptic relays. Because higher level inputs might be strongly affected by anesthetics, the late responses (as is the case with any neuronal response) might be affected by the behavioral or anesthetic state of the animal.

Another intriguing possibility is that distinct subcortical sources contribute differentially to early and late PAF responses. An important distinction has been made between areas in the ascending auditory system with so-called “specific” (lemniscal) and “diffuse” (nonlemniscal) response characteristics (Andersen et al. 1980; Calford 1983; Calford and Aitkin 1983; Imig and Morel 1983; Morel and Imig 1987; Weinberger 1993). Specific characteristics are narrow frequency tuning, short latency, and precise tonotopy, whereas diffuse characteristics are broad tuning, long latency, and imprecise tonotopy. PAF receives major projections from the “periventral nuclei,” a group of diffuse response regions surrounding the ventral division of the MGB (MGBv) (Imig and Morel 1983; Morel and Imig 1987). Imig and Morel hypothesized that the periventral nuclei are the source of broadly tuned, long latency responses in PAF; direct physiological investigations of the periventral thalamic nuclei support this hypothesis (Aitkin 1973; Phillips and Irvine 1979). PAF also receives substantial inputs from caudal MGBv, which is considered specific. While this view may be simplistic (for example, see
Rodrigues-Dagaeff et al. 1989), it is nevertheless tempting to speculate that MGBv may be the source of shorter latency, narrowly tuned PAF responses. This would not be the first example of an auditory cortical area deriving simple and complex response properties from different anatomical regions (see Fig. 5 in Rauschecker et al. 1997). This opens the possibility that anatomically distinct inputs to PAF underlie different response components with longer latency response components mediated by nonlemniscal, diffuse divisions of the ascending auditory system (Lennartz and Weinberger 1992).

Comparison with other auditory areas and functional role of PAF

How do the PAF neurons’ properties compare to those of other auditory areas? The FRAs of some neurons are reminiscent of multipeaked neurons, concentrated in A1d, but with wider frequency separation. In PAF, as in A1d, inhibitory bands were often but not always interdigitated between excitatory bands. However, many broad and diffusely responsive eFRAs in PAF (e.g., Figs. 1, R2/C5 and C6, and 3B) did not conform to the multipeaked shape observed in A1d (Sutter and Schreiner 1991). To our knowledge, broad and diffusely responsive areas have not been reported in A1. These properties are more reminiscent of response properties in the DCN than A1; however, the degree to which DCN preferentially projects to caudal MGBv or periventral MGB nuclei via the inferior colliculus (IC) is unknown. Also, anesthetic can dramatically alter DCN responses (Young and Brownell 1976), so it is unclear if the similarities between DCN and PAF would hold up in the current preparation.

The orthodox view of PAF neurons as narrow filters performing intensity discrimination is consistent with earlier data (Clarey et al. 1994; Phillips and Orman 1984), but recent data, including that from this study, reveal a more complex picture. Intensity tuning, while certainly an obvious property of PAF neurons, is just one of several properties distinguishing PAF from A1. The recent findings of this report and Heil and Irvine (1998) suggest that PAF neurons play a more complex integrative role than A1 neurons. The temporal variation in PAF response properties poses another possibility that PAF may have a temporally evolving role with later responses having a different function than earlier responses. Such a role is consistent with our results showing PAF to have more complex spectrotemporal response areas than had been originally thought.

We thank G. H. Recanzone, K. N. O’Connor, and the anonymous reviewers for helpful comments on earlier versions of the manuscript. We also thank L. Krubitzer and M. Sum for contributing histological expertise.

This investigation was supported by National Institute on Deafness and Other Communication Disorders Grant DC-02514 and by an Alfred P. Sloan Research Fellowship to M. L. Sutter. W. C. Loftus was supported by National Institute of Mental Health National Research Service Award F31 MH-11518.

REFERENCES


KITZES LM AND HOLLIGER GS. Response properties of units in the posterior auditory field deprived of input from the ipsilateral primary auditory cortex. Hear Res 100: 120–130, 1996.


