Subset of Thin Spike Cortical Neurons Preserve the Peripheral Encoding of Stimulus Onsets

Frank G. Lin1 and Robert C. Liu2,3
1The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University; 2Department of Biology, Emory University; and 3Center for Behavioral Neuroscience, Georgia State University, Atlanta, Georgia

Submitted 25 March 2010; accepted in final form 7 October 2010

Lin FG, Liu RC. Subset of thin spike cortical neurons preserve the peripheral encoding of stimulus onsets. J Neurophysiol 104: 3588–3599, 2010. First published October 13, 2010; doi:10.1152/jn.00295.2010. An important question in auditory neuroscience concerns how the neural representation of sound features changes from the periphery to the cortex. Here we focused on the encoding of sound onsets and we used a modeling approach to explore the degree to which auditory cortical neurons follow a similar envelope integration mechanism found at the auditory periphery. Our “forward” model was able to predict relatively accurately the timing of first spikes evoked by natural communication calls in the auditory cortex of awake, head-restrained mice, but only for a subset of cortical neurons. These neurons were systematically different in their encoding of the calls, exhibiting less call selectivity, shorter latency, greater precision, and more transient spiking compared with the same factors of their poorly predicted counterparts. Importantly, neurons that fell into this best-predicted group all had thin spike waveforms, suggested of suspected interneurones conveying feedforward inhibition. Indeed, our population of call-excited thin spike neurons had significantly higher spontaneous rates and larger frequency tuning bandwidths than those of thick spike neurons. Thus the fidelity of our model’s first spike predictions segregated neurons into one earlier responding subset, potentially dominated by suspected interneurons, which preserved a peripheral mechanism for encoding sound onsets and another longer latency subset that reflected higher, likely centrally constructed nonlinearities. These results therefore provide support for the hypothesis that physiologically distinct subclasses of neurons in the auditory cortex may contribute hierarchically to the representation of natural stimuli.

INTRODUCTION

The neocortex is marked not only by a stereotypy in its microcircuitry (Kozloski et al. 2001; Silberberg et al. 2002), but also by a diversity in the cell types making up these circuits (DeFelipe 1993; DeFelipe and Farinas 1992; Kawaguchi and Kubota 1997). It may not be too surprising then that a variety of spiking patterns, which may contain transient (Hromadka et al. 2008) and/or sustained (Wang et al. 2005) components, is often observed across a population of neurons in response to a single class of stimuli (Bartho et al. 2009; Chechik et al. 2006; Evans and Whitfield 1964; Galindo-Leon et al. 2009). In the auditory cortex, such different responses presumably reflect diversity in how neurons are sensitive to acoustic features within sounds (Atencio and Schreiner 2010; Sadagopan and Wang 2009) and different complex nonlinearities (Ahrens et al. 2008; Atencio et al. 2008; Sadagopan and Wang 2009) probably underlie how these neurons encode subtle variations in stimuli (Bar-Yosef and Nelken 2007; Bar-Yosef et al. 2002). Nevertheless, the degree to which diversity in acoustic sensitivity actually maps onto cortical cell type diversity or, instead, reflects an encoding strategy among presumed pyramidal cells (Chechik et al. 2006; Wang 2007) is poorly understood.

Addressing this problem requires identifying cell types from in vivo recordings in (ideally awake) animals listening to systematically varying sounds, while sampling enough neurons to quantify diversity. The latter favors extracellular methods, where a common strategy is to sort single units (SUs) based on their extracellular action potential duration (Atencio and Schreiner 2008; Bartho et al. 2004; Hromadka et al. 2008; Wu et al. 2008). Intracellular studies with morphological assessments have demonstrated that this duration correlates with various types of cortical neurons (Gray and McCormick 1996; McCormick et al. 1985; Nowak et al. 2003). For example, many thin spike SUs (those with short peak–peak durations) are thought to correspond to suspected inhibitory interneurons (Swadlow 2003), thereby providing one means to approach the cataloging of cellular diversity.

To dissect the variety in acoustic sensitivity, we started with the concept that successive stages of processing likely create more highly nonlinear neurons (Ahmed et al. 2006; Atencio et al. 2009). The corollary of this is to expect that cortical neurons would be highly nonlinear compared with auditory nerve (AN) fibers, potentially encoding acoustic features very differently. Indeed, temporal modulation is one example where the encoding is systematically transformed from the periphery to the cortex (Eggermont 2001). In contrast, sound onset encoding can actually be quite comparable between AN fibers and at least some auditory cortical neurons (Heil and Irvine 1997; Heil and Neubauer 2003; Phillips and Hall 1990). Specifically, the first spike latencies (FSLs) in both areas similarly depend on the acceleration of the amplitude envelope (Heil and Irvine 1997), likely reflecting a common mechanism for nonlinear amplitude envelope integration (Heil 2004). However, these conclusions were based on studies in anesthetized animals and a recent report comparing cortical FSLs in anesthetized versus awake conditions has demonstrated coding differences between the two (Ter-Mikaelian et al. 2007). How closely then might cortical FSLs in awake animals follow the mechanism for onset sensitivity observed at the AN and can the degree of similarity be meaningfully used to classify encoding diversity?

To answer these questions, we directly modeled the nonlinear envelope integration mechanism that can be found at the auditory periphery (Neubauer and Heil 2008) and asked how well the same mechanism, when extended to the auditory cortex, predicted its FSLs. We applied this to cortical neurons...
from awake mice listening to species-specific communication calls. Such natural vocalizations are thought to be discriminated in auditory cortical activity based on the temporal pattern of spiking (Gehr et al. 2000; Huetz et al. 2009; Liu and Schreiner 2007; Schnupp et al. 2006), thereby motivating the study of FSLs for this class of sounds. We found that the error our model produced in predicting call-evoked FSLs spanned a large range, but was not randomly distributed across SUs. There were systematic differences in sound encoding properties of neurons whose call-evoked FSLs were best versus poorly predicted by the model, presumably due to higher nonlinearities in the latter group. Most notably, the neurons whose call-evoked FSLs were best predicted tended to have thin spikes, potentially indicative of their being putative fast-spiking interneurons. Thus our data lead us to hypothesize that the neurons that best preserve the peripheral mechanism for onset encoding in the auditory cortex may actually be inhibitory rather than excitatory.

METHODS

All surgical and experimental procedures were approved by the Emory University Institutional Animal Care and Use Committee. Animals were housed under a reversed light cycle (14 h light/10 h dark) and had unrestricted access to food and water. Electrophysiology experiments were carried out during the dark cycle, corresponding to the active period for mice. We used 26 female CBA/CaJ mice, all between 14 and 24 wk old. The methods for head-post implantation, target hole surgery, lesions, and perfusion are described in detail elsewhere (Galindo-Leon et al. 2009).

Electrophysiological recordings

Electrophysiological recording locations were stereotaxically targeted by a grid of holes over the left auditory cortex, covering the ultrasound field (UF), primary auditory field (A1), and anterior auditory field (AAF) (Galindo-Leon et al. 2009; Stiebler et al. 1997). A 4- to 6-MΩ tungsten electrode (FHC, Bowdoinham, ME) was manually positioned via a micromanipulator until it was within nearly 100–200 μm of one of the target holes. The electrode was then advanced using a hydraulic microdrive (FHC), while monitoring the activity on a computer speaker. The change in background noise level that occurs when the electrode touches the cortical surface allowed us to define a reference depth. To minimize cortical dimpling, the electrode was then quickly (500 μm/s) advanced into the cortex to a depth of about 300 μm. All recordings were between 300 and 700 μm relative to the reference depth.

Single units (SUs) were isolated by adjusting the electrode depth in 1- to 5-μm steps until the SU amplitude to background noise ratio was optimized. In BrainWare software (Tucker Davis Technologies, Alachua, FL), spikes were threshold detected using both positive and negative thresholds, high-pass filtered at 300 Hz, low-pass filtered at 1 kHz, and sampled at a rate of 24 kS/s. The degree of unit isolation was based on 1) cluster analyses of waveform features in various two-dimensional spaces (first vs. second peak amplitudes, vs. peak–peak times, vs. trigger–trigger times, vs. peak–trigger times, and vs. spike area) and 2) a verification of the absence of spikes during the absolute refractory period (within 1 ms). Figure 1, A–D shows four example SUs with their respective spike waveforms and interspike intervals (ISIs). Average signal-to-noise ratio of recorded spikes (spike peak–peak normalized by the noise root-mean-square [RMS]) was 38, with standard deviation (SD) of 23 (range from 10.4 to 108 across all recordings).

We segregated SUs into thin or thick spike groups based on the peak-to-peak time of the average spike waveform (Fig. 1, A2–D2).

Acoustic stimulation

Frequency response curves were derived at 60 dB SPL by playing a set of 40 tones, with 10 ms cos2 onset and offset ramps and 60 ms long constant amplitude. Logarithmically spaced frequencies, ranging from 6.4 to 95 kHz, were presented randomly every 600 ms and repeated 15 times. The tonal rate level function of an SU at its best frequency (BF) was then derived from 30 to 70 dB SPL in increments of 10 dB SPL, repeated 10 times per amplitude in a blocked format.
SU classification and analysis

Eighteen pup calls were drawn for playback (Liu et al. 2003) from a large library of natural ultrasonic CBA/CaJ vocalizations (Fig. 2). Sound snippets were high-pass filtered in software (25 kHz corner, 8th-order Butterworth filter; butter, MATLAB, The MathWorks, Natick, MA), spectrally denoised (Liu et al. 2003), and then Hilbert transformed to extract the instantaneous frequency and amplitude envelope. These were used to resynthesize a clean version of each pup call on a silent background, multiplied by a 0.5 ms cos² onset and offset function, and scaled to a target RMS amplitude corresponding to 65 dB SPL. A maximum of 50 trials (600 ms long) of each pup call along with a blank stimulus were presented in random order, with sound onset usually beginning at 200 ms after trial onset. Occasionally, an SU drifted sufficiently in amplitude that it could no longer be isolated, in which case the call stimuli were terminated with fewer trials. Further details of the acoustic stimulation are described elsewhere (Galindo-Leon et al. 2009).

Cortical LIEFTS model for FSLs

Although different models could be implemented to predict cortical FSLs, we chose to use a “functional” peripheral model (Neubauer and Heil 2008), which has been physiologically validated for AN fiber FSLs (Heil et al. 2008) and which can be straightforwardly applied to the cortex. We preferred this over more complex subcortical models (Dugue et al. 2010; Fishbach et al. 2001, 2003), which can involve more parameters to predict cortical activity, since our purpose was to focus on how functionally similar the nonlinear envelope integration mechanism was between the periphery and cortex for encoding sound onsets.

The leaky integration, event formation, temporal summation (LIEFTS) model (Neubauer and Heil 2008) we applied accounts for a physical delay (t_{onset}) between the external stimulus and the inner hair cell response, integrates the amplitude envelope with an exponentially decaying weighting function \( P_e(t) \), and applies a biophysically plausible cubic nonlinearity (Heil and Neubauer 2003; Heil et al. 2008) to derive a stimulus-driven rate of neurotransmitter release for our purposes since different methods we implemented produced similar conclusions. Here we present results based on a binless method previously described by Chase and Young (2007), since it implements a statistical criterion that in principle accounts for the possibility that spontaneously arising first spikes can bias the latency estimate. Briefly, it collapses spike trains across trials into a list of all spikes and essentially asks whether the number of spikes occurring within a particular postonset interval exceeds what would be expected based on a Poisson distribution, with a mean rate defined by spontaneous activity and accounting for the number of trials (measured here between –150 and 0 ms with respect to stimulus onset, t_{onset}). The intervals progressively increment in time steps defined by the difference between the nth postonset spike (t_n) and our reference time (fixed here at t_{ref} = t_{onset} + 5 ms, safely accounting for the shortest possible synaptic delays to the cortex; e.g., Heil and Irvine 1997). We chose n to start at 6, comparable to that used by Chase and Young (2007). The FSL was defined as the time t_n where the probability of at least n spikes in t_n – t_{ref} could have been due to spontaneous Poisson firing first dropped below a threshold P = 0.01. Spike times after t_{onset} + 100 ms were not included, since this was beyond when even offset spiking to the longest stimuli would occur. We found our choice of parameters provided fairly consistent yet more rigorous estimates of what would be identified as the FSL “by eye” for both pure-tone and call data sets. If the spiking response to a particular pup call never satisfied the P ≤ 0.01 criterion, no statistically valid FSL was defined.

To estimate the temporal precision of first spikes across trials for each call, we used a measure analogous to the vector strength (Goldberg and Brown 1969). We mapped the period from 0 to 100 ms relative to t_{onset} to phases from 0 to 2π. Across all the trials for a particular call, each first spike was converted into a unit vector, with a phase corresponding to its latency, and these were then vector-averaged together. For example, if for all trials, the first spike occurred at 10 ms poststimulus, the length of the resultant vector for that call would be 1; if the first spikes were uniformly dispersed throughout the period, the vector would be equal to 0.

We computed the BF and tuning width for each SU based on its response to different tone frequencies at 60 dB SPL. First, a response window was defined to span the peak in an SU’s overall PSTH across all tone frequencies. The frequency yielding the maximum spike rate response was designated as the BF for that SU. To compute the bandwidth (BW) of the response curve, we determined the halfway point between the spontaneous rate and the maximum spike rate at the SU’s BF and defined the BW as the difference between the lowest and highest frequencies where the spike rate exceeded this half-max value. The tuning quality was then defined as the SU’s BF divided by the BW.

FIG. 2. Acoustic structure of pup calls. Illustration of the pup-call stimuli amplitude envelopes (top row) and their frequency spectrograms (bottom row). The calls are ordered with frequency ascending from left to right in each row and with duration increasing from bottom to top in each column.
\[ R(t') = k \left( \int_{-\infty}^{\infty} P_s(u) \exp\left(-{(t' - u)}/\tau\right) du \right)^3 \]  

The sensitivity \( k \) captures all linear proportionality factors up to this stage, \( t' \) is the time shifted by \( l_{\text{min}} \), and \( \tau \) represents the integration time constant. \( R(t') \) constitutes the hazard function of the distribution function for first spikes (Neubauer and Heil 2008) and can be used to generate the FSL probability density function (PDF)

\[ \frac{df(t')}{dt} = \left[ R(t') + R_{\text{spont}} \right] \exp\left(-\int_{-\infty}^{\infty} [R(u) + R_{\text{spont}}] du\right) \]  

where \( R_{\text{spont}} \) represents the spontaneous discharge rate, which can be estimated from the spontaneous spiking in the absence of stimuli.

To test whether this model can predict cortical first spike responses to natural calls, we increased the physical delay \( l_{\text{min}} \) and generalized the sensitivity across frequencies \( k(f) \). Intuitively, these changes presumed that the timing of first cortical spikes elicited by a call can be explained by a similar nonlinearity as observed in AN fibers, after accounting for a longer transmission distance, and a change in gain of different frequency channels. This frequency-dependent gain could arise either from the intrinsic frequency-dependent sensitivity of individual AN fibers or from the convergence of different frequency channels in the feedforward circuit to the cortex. Only responses to pure-tone stimuli were needed to build the model, which was then tested for pup call stimuli. We have implemented the cortical LIEFTS model by optimizing different combinations of essential parameters (\( l_{\text{min}}, k, \) and \( \tau \)) (Heil et al. 2008). For simplicity and because our conclusions are not essentially altered, we describe here just the results for when \( \tau \) is set to 1.4 ms (based on an estimate of the time constant of the membrane of the inner hair cell; Raybould et al. 2001) and \( l_{\text{min}} \) and \( k \) are determined as described in the following text.

We began by using the Chase and Young (2007) statistical method to find the FSLs for BF tones of varying amplitudes (Fig. 3, A and B), allowing us to construct a latency-level function (Fig. 3B, open squares). To minimize the number of parameters to fit, we set \( l_{\text{min}} \) to 0.75 \( \times \) the estimated FSL for the loudest tone at 70 dB SPL (Fig. 3B, dashed line). In a limited number of cases (e.g., SUs that were clearly nonmonotonic), we set \( l_{\text{min}} \) to 0.75 \( \times \) the shortest FSL. \( R_{\text{spont}} \) was set to the average spontaneous rate prior to the stimulus for all trials. We then solved for the value of \( k \) at BF through Eqs. 1 and 2 by using an unconstrained nonlinear optimization based on the Nelder–Mead simplex method (Fminsearch, MATLAB). We took the time at the mode of the PDF as the predicted FSL. We compared this (Fig. 3B, filled black circles) to the experimental FSL (Fig. 3B, open squares), to find the optimal value for \( k \) that minimized the RMS of the logarithm (base 10) of their ratio (root-mean-square error [RMSE]) averaged across dB SPL.

At frequencies away from a neuron’s BF, we determined \( k(f) \) by scaling the sensitivity to match FSL predictions to the derived FSLs for pure tones at different frequencies (Fig. 3C). Experimental latencies were generally longer at the highest ultrasonic frequencies, indicating that cochlear delays were not responsible for the observed spread in latencies at these frequencies. If for a specific tone frequency \( f_o \), no first spike was statistically found, we set \( k(f) \) to an SU-specific lower bound value. This was determined by when the peak probability of a predicted, acoustically driven first spike just exceeded the probability for a spontaneously generated first spike. FSLs for communication calls were then predicted (Fig. 3D) by interpolating a \( k(f) \) that matched the starting frequency of a call. Starting frequencies were used because the single-frequency calls themselves exhibited little frequency modulation (FM) near the onsets.

Although implementing the model as described earlier could be effective for predicting FSLs based on the PDF mode (see RESULTS), the predicted PDF width systematically underestimated the observed trial-by-trial variability in FSLs (e.g., Fig. 3E). This underestimation of the PDF width may be explained by the fact that the cortical

**FIG. 3.** Cortical leaky integration, event formation, temporal summation (LIEFTS) model (SUNetID 1773). A: the figure shows the spiking response to different amplitudes (30–70 dB SPL, alternating gray and white bars) at its best frequency (BF). The raster plot illustrates the first spike after stimulus onset (black diamonds) and the spikes during the spontaneous period and after the first spike (gray diamonds). The open rectangles represent the first spike latency (FSL) as defined by the Chase and Young (2007) method. B: the figure shows the FSLs at each dB SPL (open squares) and the fitted response (black circles) with the defined physical delay (\( l_{\text{min}}, \) dotted line). C: the frequency response is shown on the left with the first spikes following stimulus onset (black) and the following spikes (gray). We played pure tones and each stimulus is shown in alternating gray and white bars. We also show the corresponding linearly interpolated sensitivity across frequencies \( k(f) \). E: for call 7, we show the Gaussian-smoothed experimental first spikes (gray) and the cortical model prediction. F: the figure demonstrates the relationship between the FSL and its latency variability for all SUs (black dots; see METHODS). We found a similar power-law fit (solid gray line) to a previous cortical FSL study in anesthetized animals (Heil and Irvine 1997), shown as the dotted line.

LIEFTS model does not account for the firing rate saturation that may occur at higher stimulus amplitudes (Heil et al. 2008) and/or because temporal precision likely deteriorates after the multiple synapses leading to higher auditory stations. Indeed, consistent with previous reports (Heil and Irvine 1997; Ter-Mikaelian et al. 2007), we found a systematic increase in the FSL SD with increasing FSL. For this, we extracted a distribution of experimental first spike times for each stimulus (pure tones eliciting evoked first spikes for 50, 60, and 70 dB SPL) for SUs whose spontaneous spike rates were <5 Hz and whose poststimulus first spike time distributions significantly differed from...
spontaneous first spike times (with respect to arbitrary reference time during silence; Kolmogorov–Smirnov test, \( P < 0.05 \)). This led to a power-law relationship (Fig. 3F, solid line): \( \alpha = 0.07 \times (FSL^{-1}) \). Combining all of the preceding steps together, we arrived at a call-by-call prediction for each SU (Fig. 3D). Overall prediction quality for each SU was then measured by the relative RMSE (defined earlier) taken between predicted and experimental pup-call FSLs, but averaged only across calls with significant experimental FSLs.

**Statistical tests**

We used parametric and nonparametric statistical tests to compare group differences. Parametric tests were used only if all groups were normally distributed and did not statistically differ in their variances. Setting a significance criterion of \( P < 0.05 \), we performed the preceding tests using the two-tailed Lilliefors goodness-of-fit test for normality (lilliefors, MATLAB), and a two-tailed \( F \)-test for variance (variance, MATLAB). If both tests were statistically satisfied, we proceeded to test for group differences using the ANOVA. However, if any group failed the Lilliefors or the \( F \)-test for variance, we used nonparametric statistical tests to determine whether there were any group differences. Similar to the ANOVA, we first tested the groups using a Kruskal–Wallis one-way ANOVA (kruskalwallis, MATLAB). If either the ANOVA or the Kruskal–Wallis test returned \( P < 0.05 \), we performed group comparisons using Tukey’s honestly significant difference (HSD) criterion. To test paired data, we used a signed-test (signtest) and for correlations we used corrcoef in MATLAB, both at a \( P < 0.05 \) significance level.

**Results**

SU recordings were performed in awake, head-restrained mice. For this study, we required the SU to have an overall excitatory response to three stimulus sets, to model the timing of its evoked first spike. These included pure tones of different frequencies at 60 dB SPL, pure tones of the same frequency (BF) at different dB SPLs, and a collection of 18 ultrasound pup calls at 65 dB SPL. Only 55 of 286 SUs satisfied our restriction criterion, since many SUs could be overall inhibited (Galindo-Leon et al. 2009) or nonresponsive to one or more stimulus sets. For each SU, we fitted the cortical LIETFs model to predict its FSLs for pup calls that elicited statistically significant experimental FSLs by the Chase and Young (2007) method (see METHODS). Of the 55 SUs, 4 were further excluded from population summaries because, for them, no individual pup call evoked a statistically significant experimental FSL, even though they had an overall excitatory response.

The prediction error, characterized by the RMSE, spanned a large range equivalent to a 7.2 to 167% difference between predicted and experimental FSLs. Figure 4A, B, C shows these on a per-call (\( \times \)) as well as an averaged per-SU (solid symbols) basis for SUs segregated into RMSE quartiles: QBest contained the 25% of SUs with the lowest prediction errors for these pup calls; QPoor, 25% with the largest prediction errors; and QMid, 50% around the median. Points for both calls and SUs were close to the diagonal for the QBest group, but began deviating from it for the QMid and QPoor groups. If predicted and experimental FSLs were uncorrelated for all SUs regardless of the size of the error, then the model would simply be inappropriate for capturing the mechanisms that generate cortical first spikes (Bar-Yosef and Nelken 2007). However, we found that QBest SUs had a strong correlation between prediction and experiment (Fig. 4A, cross-correlation [cc] = 0.83, \( P < 0.001 \)) and even QPoor had a weak correlation (Fig. 4C, cc = 0.21, \( P < 0.05 \)). This indicates that the cortical LIETFs model can at least partially explain the mechanism for FSLs, albeit to varying degrees for different cortical SUs.

This result was not trivially due to better ultrasound responsiveness on the part of QBest SUs. The range of BFs was not different between groups [one-way ANOVA, \( F(2,48) = 0.47, P > 0.05 \)] and 38–44% of SUs in each group actually had BFs >50 kHz (Fig. 4D), where SUs typically showed clear responses to at least a few of the calls in the 60- to 80-kHz pup-call range. On the other hand, whether the magnitude of an SU’s prediction error was small or large also did not appear to be random. In particular, a spike waveform analysis revealed that all QBest SUs had thin spikes (circles in Fig. 4D), even though both QPoor and QMid contained an approximately equal mix of SUs with thin and thick (triangles) spikes (binomial, QMid: 11/25 thin; QPoor: 6/13 thin, \( P > 0.05 \), two-tailed). This suggests that SUs with the smallest prediction errors may be biophysically distinct from those with large errors (Hasenstaub et al. 2010).

This led us to ask whether prediction error correlated with differences in other characteristics. Indeed, QBest SUs were distinct in the nature of their nonlinear encoding of sound onsets. Figure 5 shows the rasters and pooled PSTHs to three calls (top half of each panel) and three frequency- and duration-matched pure tones (bottom half) for six SUs tuned to frequencies above (Fig. 5, A–C) or below (Fig. 5, D–F) 50 kHz, falling into either the QBest (A and D), QMid (B and E), or QPoor (C and F) groups. The two QBest examples (Fig. 5, A and D) exhibited robust onset responses for both tones and calls, with the response strength dropping as frequency increased. Because of the slight differences in mean amplitude (60 dB

---

**FIG. 4.** Prediction errors segregate SUs. A, B, C: for each group, the figure illustrates the individual predicted vs. experimental FSL (crosses) and the mean predicted FSL vs. the mean experimental FSL for each SU (circles, thin spikes; triangles, thick spikes). D: the figure demonstrates the relationship between the relative root mean square (RMS) prediction error and the SU’s BF. The dotted line represents 50 kHz and shows that the proportions of cells above and below 50 kHz are no different between the 3 groups. There were a total of 13 SUs in QBest (black symbol), 25 cells in QMid (gray symbol), and 13 cells in QPoor (open symbol). The group designation refers to whether a SU’s pup call–excited response is well predicted (QBest, top quartile), poorly predicted (QPoor, bottom quartile), or within the middle 50% (QMid) of prediction errors across the population.

---

*J Neurophysiol.* • VOL 104 • DECEMBER 2010 • www.jn.org
SPL for pure tones vs. 65 dB SPL for calls) and amplitude onsets (10 ms cos² ramp for pure tones vs. rapid onsets for calls; Fig. 2), call and tone responses were not always identical (compare call 17 to the 76-kHz tone response for SUnitID 1773). Importantly though, the nonlinearities in the cortical LIEFTS model correctly accounted for these acoustic differences.

In the case of QMid and QPoor, though, some calls elicited responses that were completely unexpected based on the tone response, even though the calls were single-frequency whistles. For example, SUnitID 1363 (Fig. 5C) responded transiently to tone onsets, but fired robustly near call offsets. Similarly, the response of QMid SUnitID 1355 (Fig. 5E) to tones was imprecise, but was well-locked to sound onset for calls. On the other hand, QPoor SUnitID 1594 (Fig. 5F) fired transiently and reliably for tones, but was unresponsive to calls 13, 15, and 17. We also observed call-evoked onset inhibition in some SUs, such as QMid SUnitID 1745 in response to call 17 (Fig. 5B). Such inhibition was not accounted for by our purely excitatory model. Thus the failure of the cortical LIEFTS model for QMid and QPoor SUs coincided with cases where the response to pure tones could not be extrapolated to explain actual call-evoked responses, suggesting more complex sensitivities to the acoustics of a sound’s onset.

Indeed, sensitivity to onset frequency was noticeably more consistent for QBest compared with QPoor SUs, based on the statistically estimated FSL for individual natural calls. QBest and QMid SUs had clearly delayed FSLs for higher- versus lower-frequency calls (Fig. 6A, black and gray diamonds, signed-test, zQB = 5.1, NQB = 39, P < 0.001, zQM = 3.9, NQM = 48, P < 0.001), whereas FSLs for QPoor SUs were not systemati-
cally different (open diamonds, signed-test, $z_{QP} = 0.6$, $N_{QP} = 11$, $P > 0.05$). Similarly, QBest and QMid SUs showed significantly later FSLs for slower amplitude onsets compared with faster onsets (Fig. 6B, black and gray diamonds, respectively, mostly above the diagonal, signed-test, $z_{QP} = 2.8$, $N_{QP} = 19$, $P < 0.01$, $z_{QM} = 3.5$, $N_{QM} = 2$, $P < 0.001$). However, FSLs for QPoor SUs across their populations were not significantly different for the two types of sounds (open diamonds, signed-test, $z_{QP} = 0$, $N_{QP} = 5$, $P > 0.05$), suggesting that the shape of the envelope onset of the ultrasound was not a reliable driver of their first spikes.

Figure 6 also suggests that the absolute latencies of pup-excited responses may be systematically different between groups. QBest FSLs were generally earlier than QPoor FSLs on a per-call basis [Fig. 7A, Kruskal–Wallis, $\chi^2(2,509) = 135.3$, $P < 0.001$, Tukey’s post hoc: QBest–QMid, $P < 0.05$; QBest–QPoor, $P < 0.05$; QMid–QPoor, $P < 0.05$], and averaged per SU [one-way ANOVA, $F(2,48) = 10.7$, $P < 0.001$, Tukey’s post hoc: QBest–QMid, $P < 0.05$; QBest–QPoor, $P < 0.05$; QMid–QPoor, $P < 0.05$]. Moreover, first spikes were also more precisely timed trial by trial (quantified by vector strength; see METHODS) for QBest compared with QPoor on a per-call basis [Fig. 7B, Kruskal–Wallis, $\chi^2(2,509) = 79.3$, $P < 0.001$, Tukey’s post hoc: QBest–QMid, $P < 0.05$; QBest–QPoor, $P < 0.05$; QMid–QPoor, $P < 0.05$], as well as averaged per SU [one-way ANOVA, $F(2,48) = 3.2$, $P < 0.05$, Tukey’s post hoc: QBest–QMid, n.s.; QBest–QPoor, $P < 0.05$; QMid–QPoor, n.s.].

Another difference between the best- and worst-predicted SUs arose in the selectivity for individual calls. QBest SUs generally exhibited statistically significant FSLs to far more of the calls than QPoor SUs [Fig. 7C, Kruskal–Wallis, $\chi^2(2,48) = 7.3$, $P < 0.05$, Tukey’s post hoc: QBest–QMid, n.s.; QBest–QPoor, $P < 0.05$; QMid–QPoor, n.s.]. This result could not be explained simply by a difference in spontaneous firing between groups [Fig. 7D, Kruskal–Wallis, $\chi^2(2,48) = 4.1$, $P > 0.05$].

Strikingly, FSL prediction error also correlated with systematic differences in spiking after the first spike. This is evident from comparing the call-evoked PSTHs for the three groups (Fig. 8A, population-averaged and pooled over all 18 calls, normalized by spontaneous rate and smoothed). QBest SUs (thick black line) consistently showed strong transient firing at sound onset. Some SUs, such as SUni1ID 1773 (Fig. 5D), also produced sustained firing, but this was not that common. On the other hand, the averaged discharge pattern for QPoor SUs (thin dashed line) was somewhat delayed, lacked a strong onset, and instead held a more stable level of firing during and
beyond the stimulus. Finally, QMid SUs exhibited a more heterogeneous group of responses that seemed to exhibit response characteristics between QBest and QPoor. These data indicate that SU groups differed not only in terms of how they fired first spikes, which was the basis for classifying them, but also in terms of how they fired subsequent spikes.

The population differences were apparent even at the individual SU level. Evaluating call-pooled, normalized PSTHs on an SU-by-SU basis, QBest SUs (Fig. 8B, black symbols) had significantly shorter half-max PSTH durations [right panel; Kruskal–Wallis, χ²(2,48) = 9.3, P < 0.01, Tukey’s post hoc: QBest–QMid, P < 0.05; QBest–QPoor, P < 0.05; QMid–QPoor, n.s.] and earlier half-max PSTH latencies [bottom panel; Kruskal–Wallis, χ²(2,48) = 15.9, P < 0.01, Tukey’s post hoc: QBest–QMid, n.s.; QBest–QPoor, P < 0.05; QMid–QPoor, P < 0.05] compared with QPoor. The majority of QBest SUs therefore clustered tightly toward the bottom left corner of Fig. 8B, where transient onset SUs should lie, whereas QPoor SUs were generally found more to the right and top portions of Fig. 8B, where offset and more sustained SUs should lie. QMid SUs were scattered throughout this plane and were not different from QPoor SUs in PSTH duration, but were for PSTH latencies. As a result, having a transient onset response to calls was not sufficient for an SU to belong to QBest (gray and black symbols in bottom left quadrant of Fig. 8B, <20 ms latency and <20 ms duration), although it was an important factor (11/13 black symbols are in that quadrant).

A summary of direct comparisons between QBest and QPoor properties is shown in Table 1, Box 1, reiterating the systematic call encoding differences that emerged simply by grouping SUs according to their relative FSL prediction error under the cortical LIFETS model. By using relative error, we did not artificially bias our results toward SUs with earlier, less variable FSLs (Fig. 3F). Indeed, two QBest SUs in Fig. 8B fell outside of the onset-transient group because they had noticeably later and longer responses. Thus it is interesting to find that the majority of QBest SUs nevertheless were among the earliest firing in response to calls.

This raises the question of whether absolute timing error might be more pertinent for segregating SUs. To address this, we regrouped SUs into quartiles based on the RMS absolute FSL difference between prediction and experiment (Table 1, Box 2), keeping all fits and predictions the same. Importantly, we reached the same overall conclusions concerning differences between properties of QBest and QPoor SUs, in some cases at even greater levels of significance. Moreover, 9 (10) SUs remained in the lowest (highest) error quartile regardless of whether relative or absolute measures were used (Table 1, Box 3). All differences between these intersected groups of QBest and QPoor SUs remained significant. These results provide confidence that one subgroup of short latency, precisely firing, mostly transient cortical SUs likely encode the acoustic onsets of these calls in a similar way as at the auditory periphery.

Finally, since the QBest SUs all had thin spikes, we evaluated criteria that might indicate whether thin spike SUs correspond to suspected fast-spiking interneurons (Table 2). Ignoring error grouping, we found that our thin spike SUs had significantly higher spontaneous firing rates and were recorded at a shallower depth than thick spike SUs. Moreover, the bandwidth of the tonal frequency response curves was significantly larger for the former (smaller quality factor, defined by BF/BW; see METHODS). In response to pup calls, FSLs were also significantly earlier for thin spike SUs. These results suggest that as a population, thin spike SUs have characteristics consistent with being possible fast-spiking interneurons, although we cannot definitively know this for any one SU without intracortical recordings. Importantly though, the differences between QBest and QPoor SUs was not just due to a change in the balance between thin and thick spike SUs. Restricting only to thin spike SUs in both groups, we found that the FSL and FSL precision per SU and per call as well as the PSTH half-max were still significantly different, even though spontaneous activity and

### Table 1. Using relative or absolute error segregates a similar population of QBest and QPoor

<table>
<thead>
<tr>
<th>Factor</th>
<th>Segregated by Relative Error</th>
<th>Segregated by Absolute Error</th>
<th>Intersected SUs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QBest</td>
<td>QPoor</td>
<td>QBest</td>
</tr>
<tr>
<td>Number of SUs</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Number of call responses</td>
<td>12.2 ± 1.6 *</td>
<td>6.8 ± 0.8 **</td>
<td>13.6 ± 1.4 **</td>
</tr>
<tr>
<td>FSL per call, ms</td>
<td>18.0 ± 0.8 ***</td>
<td>48.0 ± 2.4 ***</td>
<td>14.4 ± 0.4 ***</td>
</tr>
<tr>
<td>FSL per SU, ms</td>
<td>20.1 ± 2.6 *</td>
<td>46.2 ± 4.6 **</td>
<td>14.7 ± 0.9 **</td>
</tr>
<tr>
<td>FSL precision per call</td>
<td>0.77 ± 0.02 ***</td>
<td>0.53 ± 0.01 **</td>
<td>0.79 ± 0.05 ***</td>
</tr>
<tr>
<td>FSL precision per SU</td>
<td>0.71 ± 0.05 *</td>
<td>0.53 ± 0.06 **</td>
<td>0.75 ± 0.04 *</td>
</tr>
<tr>
<td>Call PSTH half-max, ms</td>
<td>13.0 ± 1.7 ***</td>
<td>43.7 ± 7.6 ***</td>
<td>9.6 ± 0.7 ***</td>
</tr>
<tr>
<td>Call PSTH duration, ms</td>
<td>15.4 ± 1.8 **</td>
<td>35.0 ± 6.5 **</td>
<td>12.2 ± 0.6 **</td>
</tr>
<tr>
<td>% thin spike SUs</td>
<td>100 *</td>
<td>46 *</td>
<td>100 ***</td>
</tr>
<tr>
<td>Spontaneous rate, spikes/s</td>
<td>11.8 ± 3.7 n.s.</td>
<td>5.5 ± 2.0 **</td>
<td>14.3 ± 3.7 ***</td>
</tr>
<tr>
<td>Tuning width, BF/BW</td>
<td>1.7 ± 1.6 *</td>
<td>2.3 ± 0.2</td>
<td>1.4 ± 0.2 **</td>
</tr>
</tbody>
</table>

Values are means ± SE for each of the analyses that were done previously for groups segregated using RMS relative error (columns 2 and 3, box 1) and RMS absolute error (columns 4 and 5, box 2). Columns 6 and 7 (Box 3) represent those SUs that were segregated by both methods. For significance: *P < 0.05; **P < 0.01; ***P < 0.001; n.s., not significant.

### Table 2. Thin and thick spiking SUs have different spiking characteristics

<table>
<thead>
<tr>
<th>Factor</th>
<th>Thin Spikes</th>
<th>Thick Spikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of SUs</td>
<td>30</td>
<td>21</td>
</tr>
<tr>
<td>Spontaneous rate, spikes/s</td>
<td>10.6 ± 2.0</td>
<td>*** 2.2 ± 0.4</td>
</tr>
<tr>
<td>Tuning width, BF/BW</td>
<td>1.9 ± 0.1</td>
<td>*** 2.7 ± 0.1</td>
</tr>
<tr>
<td>Depth rel. surface, μm</td>
<td>484.1 ± 20.5</td>
<td>553.4 ± 18.8</td>
</tr>
<tr>
<td>FSL per call, ms</td>
<td>22.6 ± 0.8</td>
<td>*** 41.4 ± 1.6</td>
</tr>
<tr>
<td>FSL per SU, ms</td>
<td>24.3 ± 2.3</td>
<td>*** 45.0 ± 3.2</td>
</tr>
</tbody>
</table>

The table shows the number of thin and thick spiking SUs. There were significant differences between the two groups of SUs in the spontaneous rate, tuning width, recording depth, FSL per call, and FSL per SU.
tuning bandwidth were not (presumably since all SUs had thin spikes). Although we did not find a significant difference in the number of call responses, a trend similar to that found in Table 1 was present ($t$-test, $t = 2.0$, df = 17, $P = 0.06$; $Q_{\text{Best}}$: 12.2 ± 1.6; $Q_{\text{Poor}}$: 7.2 ± 1.2). Thus the degree to which an SU’s onset response to pup calls follows the cortical LIEFTS model appears to segregate out a distinct subpopulation of thin spiking neurons, some of which may be putative fast-spiking inhibitory interneurons.

**DISCUSSION**

A main conclusion of this work was that diversity in auditory cortical responses to sound onsets can be meaningfully sorted according to how well the first sound-evoked spikes are explained by the nonlinear amplitude envelope integration mechanism present at the periphery. Using data collected from awake mice, we arrived at this by applying the LIEFTS model of sound pressure integration to predict cortical responses (Fig. 3). Cortical SUs whose responses to ultrasonic calls (Fig. 2) were best predicted by this model were systematically different from poorly predicted SUs (Fig. 4): the former responded to more calls with earlier and more precise transient onsets (Figs. 7 and 8) and had greater sensitivity to the acoustic features of natural call onsets (Fig. 6). Most strikingly, these SUs, accounting for roughly 18–25% of pup-call–excited SUs, all had thin extracellular spike waveforms, suggesting that they could be physiologically distinct.

In contrast, the worst predicted SUs, consisting of a mix of thin and thick spike SUs, were generally more call-selective, fired later, and less precisely (Figs. 7 and 8) and were not consistently sensitive to sound frequency and amplitude envelope (Fig. 6). However, the fact that their first spikes did not closely follow the peripheral mechanism for responding to sound onsets did not imply that these SUs were acoustically unresponsive, since all showed an excitation to the calls as a whole. Instead, their activity potentially resulted from higher nonlinearities created de novo by processing within the central auditory system. Thus our results suggest that hierarchical processing by the auditory system may preserve a peripheral encoding of stimulus onsets within some short-latency, thin spike cortical neurons and construct a more complex acoustic selectivity within other neurons.

**Model caveats**

First, we used a functional rather than a detailed model of subcortical circuitry (Dugue et al. 2010), incorporating only effective excitatory input across frequency channels. Inhibition, although not explicitly included, could have contributed so long as it did not abolish the excitatory response and might explain why some observed latencies were longer than predicted (Fig. 4C). Moreover, interactions between frequency channels simultaneously excited by a sound were ignored, since we assumed that only the channel centered at a sound’s initial frequency contributed to the cortical neuron’s acoustic sensitivities. Our particular class of ultrasonic calls contained single frequencies with little frequency modulation (FM), so this “winner-take-all” approach ensured that the frequency channel that should respond best had the strongest influence. Indeed, an earlier multiunit study in anesthetized mice also using pup calls corroborated that auditory cortical spiking on average does not detect the slight FM in these calls (Liu et al. 2006). Nevertheless, although our assumption appeared valid for our best-predicted SUs, the poorer predictions of some other SUs may have arisen from sensitivity to a call’s precise frequency trajectory. Importantly though, this would still mark a distinction between how the acoustics of sound onsets drive spiking in $Q_{\text{Poor}}$ versus $Q_{\text{Best}}$ SUs.

Second, our model was relatively simplistic, but had the advantage of only needing to optimize the parameter $k$ to be able to predict FSLs that were significantly correlated with experimental FSLs for all SU groups. In implementing the LIEFTS model, we have also tried using a multistep fitting procedure (Heil et al. 2007) to simultaneously optimize either $k$ and $L_{\text{min}}$ or $k$, $L_{\text{min}}$ and $\tau$. Although more complex, these cases nevertheless produced comparable differences in the physiological properties of segregated SUs, thus providing some assurance that our conclusions were not highly sensitive to the details of model implementation.

Third, although incorporating a frequency-dependent $k(f)$ accounted for the FSL of $Q_{\text{Best}}$ SUs to the ultrasonic whistles, its validity for predicting either subsequent spikes or more complex sound (e.g., noise or multiharmonic calls) responses was not directly evaluated. In principle, the LIEFTS algorithm can model a hair cell neurotransmitter release rate throughout a sound and can be combined with a cochlear filter bank (Fishbach et al. 2001, 2003). Combinations of central excitation and inhibition could then be fitted, creating cortical discharge patterns that could be matched to data. Importantly though, these steps were unnecessary to draw the conclusions reached here about the distinct, stereotyped properties of $Q_{\text{Best}}$ SUs.

**Classification of thin and thick spike SUs**

Since our $Q_{\text{Best}}$ SUs all had thin spike waveforms, it is tempting to speculate that these map onto a functionally distinct subset of cortical neurons. The idea that cortical cell types can be classified at least partly by extracellular spike durations has support. Theoretically, thin spikes may reflect neuronal expression of particular variants of potassium and sodium voltage-gated ion channels that help minimize energy consumption during rapid spiking (Hasenstaub et al. 2010). Experimentally, electrophysiological and morphological traits of thin and thick spike neurons recorded intracellularly are clearly different. The latter are often regular-spiking pyramidal or spiny stellate neurons that have long afterhyperpolarizations, adapt strongly to depolarizing currents, and are found across cortical layers 2–6 (McCormick et al. 1985). However, regular-spiking nonpyramidal neurons can also have thicker spikes (Kawaguchi and Kubota 1993, 1997), indicating that some interneurons may be present along with pyramidal cells in a population of thick spike SUs. Neurons with thin spikes can correspond to either chattering pyramidal cells or fast-spiking GABAergic interneurons, both of which tend to lie more in supragranular and granular cortical layers (Azouz et al. 1997; Gray and McCormick 1996; Nowak et al. 2003). Between them, fast-spiking neurons fire more rapidly, with very little spike-frequency adaptation during depolarizing current injection (McCormick et al. 1985). Chattering neurons produce bursts that give rise to a clearly bimodal ISI distribution, during
both current injection (Nowak et al. 2003) and long-duration sensory stimulation (Gray and McCormick 1996).

Clearly, then, intracellular recordings with morphological identifications are needed to unequivocally identify thin spike neurons as fast-spiking GABAergic interneurons, even though this is often assumed to hold for most thin spiking SUs (Atencio and Schreiner 2008; Swadlow 2003). Thus some degree of error is possible in inferring that our thin spike SUs may be mostly fast-spiking interneurons, especially given the comparatively frequent rate at which we encountered thin spike SUs without directly targeting them (compare Atencio and Schreiner 2008 and Rose and Metherate 2005 with Wu et al. 2008). In fact, in a few instances (4 of 34 thin spike SUs), we did observe bimodal ISI distributions of the type previously associated with chattering cells (Gray and McCormick 1996; Nowak et al. 2003), including the SU illustrated in Fig. 1B. However, since ISI parameters can depend on the nature of the stimulation (Azouz et al. 1997), and our sounds were themselves very brief, this should not be taken as definitive evidence for chattering neurons.

In favor of the possibility that our thin spike population might nevertheless be dominated by suspected fast-spiking interneurons, we found that the average depth of the recording sites for thin spike SUs was significantly shallower than that for thick spike SUs (Table 2), consistent with the expected spatial distribution described earlier. Moreover, the characteristics of our thin spike SU population (Table 2) were in agreement with other reports about extracellularly identified “suspected interneurons” (Atencio and Schreiner 2008; Bruno and Simons 2002; Swadlow et al. 1998): higher spontaneous rate, broader stimulus tuning (to frequency in our case), and earlier latencies relative to the thick spike SUs. Note that these differences remained significant even if the four SUs with bimodal ISIs were removed from the comparisons in Table 2 (data not shown). Furthermore, although three of these SUs fell into the QBest group by relative error segregation, these were the longest-latency SUs in that group (three black circles in center of Fig. 8B). Since none was among the intersected QBest group, where differences between QBest and QPoor SUs actually became more significant, this further supports the possibility that most of our best-predicted SUs form one physiologically distinct group that may correspond to suspected fast-spiking interneurons.

If the foregoing were true, then we should ask why neurons that more faithfully preserve the peripheral nonlinear mechanism for sound onset encoding might convey inhibition rather than excitation in the cortex. Speculatively, one answer may be simply that purely excitatory scaling with the integrated envelope could be too metabolically costly to encode sounds (Hasenstaub et al. 2010) and might also saturate the dynamic range of neuronal spiking. Indeed, sparse cortical coding has been hypothesized to be enabled by precisely timed inhibition (Wolfe et al. 2010). More abstractly, it may be that the functional role of pyramidal neurons in auditory cortex is not simply to represent acoustic features, but to extract so-called auditory objects (Ulanovsky et al. 2004). In this case, perhaps the activity of “acoustically faithful” feedforward inhibition might serve to synchronize different pyramidal neurons encoding higher-level features of auditory objects (Bush and Selnoski 1996).

Relation to prior work on first spike coding

First, our analysis of first spikes contributes to a growing literature on this particular feature of the spiking code. There is evidence from the visual (Oram and Perrett 1992; Thorpe et al. 1996) and somatosensory (Johansson and Bizziennks 2004) systems, suggesting that these spikes carry perceptually and behaviorally relevant information for fast sensory decisions (Thorpe et al. 2001). For audition, the cortical LIEFTS model essentially converted the strength of an acoustic stimulus into a first spike timing delay by trading quieter sound pressures for longer latencies in QBest SUs (Fig. 3, A and B). Thorpe (1990) proposed exactly this kind of “analog-to-delay” converter as a basis for a “rank-order” spike-timing–based population code (Thorpe et al. 2001). Notably, to decode such a representation, a feedforward shunting inhibitory circuit pooling across converging inputs has been suggested (Thorpe 1990; Thorpe et al. 2001), a role potentially filled by the best-predicted, putative fast-spiking SUs.

Second, our FSL data for natural calls complemented earlier studies of cortical FSL that used simpler synthetic stimuli (pure or frequency-modulated tones and noise bursts), carried out mainly in anesthetized preparations (Fishbach et al. 2001; Furukawa and Middlebrooks 2002; Heil 1997; Heil and Irvine 1997; Nelken and Versnel 2000; Phillips 1998; Phillips and Hall 1990). Although first spikes were generally found to be early and precise in those cases, Ter-Mikaelian et al. (2007) recently concluded that cortical FSLs for BF tones become later and less precise in awake animals (however, for an example in which overall temporal resolution for discriminating certain sounds can become finer, see Huetz et al. 2009). Our data (Fig. 3F) were consistent with those reported by Ter-Mikaelian (2007) since our FSL SDs were quantitatively similar to their awake cortical data. Interestingly, they found a few neurons (roughly 9 of 54 = 17%) with short latencies <20 ms and high precision <2 ms. Their population PSTH for sinusoidally amplitude-modulated tones also revealed a strong, distinct spiking contribution in the awake A1 from short-latency cells. Our results suggest that rather than simply representing variability in cortical responses, these neurons may have corresponded to a systematically separable subpopulation of cortical neurons analogous to our QBest SUs.

Finally, our work extended prior cortical modeling efforts that used selected data collected from anesthetized animals (Fishbach et al. 2001, 2003; Heil 1997, 2004). In contrast to their focus on responses with prominent onsets, we did not preselect response types, requiring only an excitatory response from which we could extract first spikes. This led us to conclude that most cortical SUs in awake animals actually have responses to sound onsets that are not well described by the nonlinear envelope integration mechanism at work in auditory nerve fiber firing (Heil et al. 2008; Neubauer and Heil 2008). This could be considered either trivial because of the high number of intervening processing stages, in which case the success for the best-predicted SUs is surprising, or unexpected because of this mechanism’s success in anesthetized animals (Heil and Irvine 1997; Phillips and Hall 1990).

Implications for hierarchical processing

The fact that QBest SUs in the auditory cortex preserve the peripheral mechanism for sound onset integration suggests that
neurons in earlier auditory stations should show a similar encoding of sound onsets. In particular, the presence of strong transient responses has been documented subcortically in both anesthetized (Blackburn and Sachs 1989; Pfeiffer 1966; Rhode and Smith 1986a,b; Syka et al. 2000; Woolley and Cassidy 2004; Zheng and Escabi 2008) and nonanesthetized preparations (Pollak et al. 1978; Shofner and Young 1985). However, classifying neurons solely by their onset PSTH may not be equivalent to identifying QBest neurons because discharge patterns can change depending on acoustic stimuli (Pfeiffer 1966; Pollak et al. 1978; Wang et al. 2005). In fact, we found onset components to call responses even among non-QBest SUs (Fig. 8B), as well as sustained components for some stimuli among QBest SUs (e.g., Figs. 3C and 5D). Thus future studies could benefit from classifying auditory neurons by their expected (i.e., modeled) transformation of sounds instead of their stimulus-specific PSTH. This may provide a more consistent taxonomy that reflects the neural sensitivities to acoustic features such as sound amplitude.

Functionally, QBest and QPoor SUs differed in their sound encoding and in principle might be distinct neuronal groups at the same level of cortical processing (e.g., different pyramidal cells). In particular, QBest SUs may preserve the peripheral mechanism to sound onsets, whereas QPoor SUs may represent subsequent acoustic modulations or nonauditory information. This difference could relate to the stimulus-synchronized and nonstimulus-synchronized neurons reported in the awake hamster auditory cortex (Lu et al. 2001). However, results from our spike waveform analysis argue strongly against a parallel mechanism to sound onsets, whereas QBest SUs are all thin spiking neurons with earlier and more transient first spike responses, they may instead be a hierarchically different class of neurons from those in the later-firing, more heterogeneous QMid and QPoor groups.

ACKNOWLEDGMENTS

We thank B. Kocher, Y. Zhang, and E. Galindo-Leon for electrophysiology assistance and T. Ivanova for mouse husbandry.

GRANTS

This work was supported by National Institute on Deafness and Other Communication Disorders Grant DC-008343, National Science Foundation Science and Technology Center for Behavioral Neuroscience Grant IBN-9876754, and Integrative Graduate Education and Research Traineeship Program Grant DGE-0333411.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

REFERENCES


