BOLD fMRI study of ultrahigh frequency encoding in the inferior colliculus

Patrick P. Gao, Jevin W. Zhang, Russell W. Chan, Alex T.L. Leong, Ed X. Wu

Laboratory of Biomedical Imaging and Signal Processing, The University of Hong Kong, Pokfulam, Hong Kong SAR, China
Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China
Department of Anatomy, The University of Hong Kong, Pokfulam, Hong Kong SAR, China
Department of Medicine, The University of Hong Kong, Pokfulam, Hong Kong SAR, China

ABSTRACT

Many vertebrates communicate with ultrahigh frequency (UHF) vocalizations to limit auditory detection by predators. The mechanisms underlying the neural encoding of such UHF sounds may provide important insights for understanding neural processing of other complex sounds (e.g., human speeches). In the auditory system, sound frequency is normally encoded topographically as tonotopy, which, however, contains very limited representation of UHFs in many species. Instead, electrophysiological studies suggested that two neural mechanisms, both exploiting the interactions between frequencies, may contribute to UHF processing. Neurons can exhibit excitatory or inhibitory responses to a tone when another UHF tone is presented simultaneously (combination sensitivity). They can also respond to such stimulation if they are tuned to the frequency of the cochlear-generated distortion products of the two tones, e.g., their difference frequency (cochlear distortion). Both mechanisms are present in an early station of the auditory pathway, the midbrain inferior colliculus (IC). Currently, it is unclear how prevalent the two mechanisms are and how they are functionally integrated in encoding UHFs. This study investigated these issues with large-view BOLD fMRI in rat auditory system, particularly the IC. UHF vocalizations (above 40 kHz), but not pure tones at similar frequencies (45, 55, 65, 75 kHz), evoked robust BOLD responses in multiple auditory nuclei, including the IC, reinforcing the sensitivity of the auditory system to UHFs despite limited representation in tonotopy. Furthermore, BOLD responses were detected in the IC when a pair of UHF pure tones was presented simultaneously (45 & 55 kHz, 55 & 65 kHz, 45 & 65 kHz, 45 & 75 kHz). For all four pairs, a cluster of voxels in the ventromedial side always showed the strongest responses, displaying combination sensitivity. Meanwhile, voxels in the dorsolateral side that showed strongest secondary responses to each pair of UHF pure tones also showed the strongest responses to a pure tone at their difference frequency, suggesting that they are sensitive to cochlear distortion. These BOLD fMRI results indicated that combination sensitivity and cochlear distortion are employed by large but spatially distinctive neuron populations in the IC to represent UHFs. Our imaging findings provided insights for understanding sound feature encoding in the early stage of the auditory pathway.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Many vertebrates have evolved the ability to communicate with ultrahigh frequency (UHF) vocalizations to decrease the likelihood of detection by predators (Brudzynski, 2009, 2013; Gadziola et al., 2012; Portfors, 2007; Takahashi et al., 2010). For example, male mice emit vocalizations with frequencies up to 110 kHz when encountering female mice, and female mice produce frequency-modulated upsweeps above 60 kHz when their pups are removed from the nest (Portfors, 2007). Rats emit short chirping 50 kHz vocalizations during juvenile play, male agonistic behaviors and sexual behaviors (Brudzynski, 2009; Portfors, 2007). Recently, there is increasing interest in studying the neural mechanisms underlying the learning, production and representation of such vocalizations (Arriaga and Jarvis, 2013; Woolley and Portfors, 2013) because it may improve our understanding of complex sound (e.g., speeches) processing in humans (Rauschecker and Scott, 2009).

One important question is how the auditory system encodes UHFs, in order to process vocalizations that contain them (Romanski and Averbeck, 2009; Woolley and Portfors, 2013). In the auditory system, the majority of neurons are sensitive to only a narrow spectrum of sounds centered about a characteristic frequency (CF). The topographic alignment of the CFs (i.e., the tonotopy) is a fundamental organizing principle maintained from the cochlea to the auditory cortex (Ehret and Fischer,
1991; Malmierca, 2003). For example, in the primary auditory nucleus in the midbrain, called the inferior colliculus (IC), which is a compulsory relay for all ascending auditory projections (Malmierca, 2003) and a region that efficiently encodes vocalizations (Holmstrom et al., 2010), the tonotopic organization is a basic feature in its central nucleus (CNC), where the CFs of neurons run from low to high along the dorsolateral-ventromedial dimension (Cheung et al., 2012a, 2012b; De Martino et al., 2013; Malmierca et al., 2008; Ress and Chandrasekaran, 2013; Schreiner and Langner, 1997; Yu et al., 2005). In the cortical regions of the IC, the frequency organization is usually non-tonotopic, with neurons showing complex frequency tuning properties, such as broadband or multiplex receptive fields (Duque et al., 2012; Hernandez et al., 2005; Malmierca et al., 2011; Stebbings et al., 2014). Given the importance of UHFs in social communications, they are expected to occupy a substantial area in the tonotopic map of the CNC, or the frequency receptive fields of neurons in non-tonotopic IC regions. But surprisingly, there is limited representation of UHFs in the auditory system of many species. In mouse IC, neurons with CF higher than 60 kHz are rarely found and occupy a very small tonotopic area compared to those with lower CFs (Stiebler and Ehret, 1985; Yu et al., 2005). Similarly, in rat IC, representations of frequencies higher than 40 kHz are seldom delineated (Cheung et al., 2012a, 2012b; Malmierca et al., 2008). This mismatch immediately poses a basic question: how UHFs are encoded in the auditory system of these species.

Current evidence of UHF encoding is found in electrophysiological studies that probed neuronal responses to behaviorally relevant sound in the auditory system. In the mustached bat, neurons in the auditory cortex as well as the IC were found to display excitatory or inhibitory responses when two tones with different frequencies were presented simultaneously (two-tone paradigm) with one of them in the frequency range of bat biosonar or vocalizations (Leroy and Wenstrup, 2000; Mittmann and Wenstrup, 1995; Suga et al., 1979). This neuron property, i.e., responding best to the combination of two different frequency components, was called combination sensitivity (Leroy and Wenstrup, 2000; Portfors and Felix, 2005; Portfors and Wenstrup, 2002). Combination sensitivity was believed to allow neural representation of spectral and temporal information conveyed in echolocation signals or social vocalizations, and was suggested to be a common mechanism for processing complex sound (Portfors, 2004; Portfors and Wenstrup, 2002). Combination sensitive neurons were also found in mouse IC (Felix and Portfors, 2007; Portfors and Felix, 2005). In both bat and mouse, the facilitatory or inhibitory frequencies were often found to be in the UHF range (Portfors and Felix, 2005; Portfors and Wenstrup, 2002), indicating that combination sensitivity might be a neural mechanism underlying UHF encoding shared by different mammals.

Recently, it was observed that neurons in the central auditory system could also respond to the two-tone paradigm by exploiting cochlear nonlinearities. When simultaneously stimulated with two different frequencies (f1 and f2, f2 > f1), the cochlear characterized by mechanical nonlinearity will generate distortion tones with frequencies at e.g. f2−f1 (quadratic) (Robles et al., 1991). The distortion products can propagate along the cochlea to their CF site on the basilar membrane and then excite neurons with the same CF at higher levels of the auditory pathway, including the cochlear nucleus (Faulstich and Rossil, 1999), the IC (Abel and Rossil, 2009; McAlpine, 2004) and the auditory cortex (Purcell et al., 2007). It has been shown that at least in some cases, cochlear distortion products were responsible for the neuronal activities that were thought to arise from combination sensitivity (McAlpine, 2004). It was also noticed that some IC neurons could respond to the cochlear distortion of UHF tones (Portfors and Roberts, 2014; Portfors et al., 2009), suggesting that cochlear distortion could also be a mechanism underlying UHF encoding.

By demonstrating the neuronal responses in multiple auditory nuclei that arise from combination sensitivity and cochlear distortion, the above-mentioned studies provided important clues for understanding UHF encoding. However, it remains unclear how prevalent these mechanisms are in an auditory nucleus, since relatively small numbers of neurons were usually recorded in those electrophysiological studies. Furthermore, the previous studies were limited by the large variation of frequency values used in the two-tone paradigm. Since different neurons display different tuning properties, one of the two frequencies was usually fixed at the CF or limited in the frequency range of a vocalization that the neuron was sensitive to (Portfors and Felix, 2005; Portfors et al., 2009; Portfors and Wenstrup, 2002). This strategy could effectively reduce the large number of combinations to test, but also induce large variation in the frequency values tested in a study. With such limitations, it was not addressed whether UHFs are encoded in a generally organized way in the auditory system as like the tonotopic encoding of lower frequencies is, or only particular UHFs can be encoded by some specialized neurons. Moreover, it remains unknown how the two mechanisms are combined by the auditory system in processing UHFs, e.g. whether they are adopted by spatially different neuron populations or the same set of neurons at different temporal stages.

Blood-oxygen-level-dependent (BOLD) functional magnetic resonance imaging (fMRI) (Ogawa et al., 1990) with large field of view is a non-invasive technique that can measure the hemodynamic responses as neural correlates (Logothetis et al., 2001) in multiple brain structures with relatively high spatial and temporal resolutions. Recent development in BOLD fMRI has been demonstrated for studying the rodent auditory system (Cheung et al., 2012a, 2012b; Gao et al., 2014; Lau et al., 2013, 2015; Yu et al., 2009; Zhang et al., 2013), along with previous auditory fMRI studies conducted on humans (Barton et al., 2012; De Martino et al., 2013; Sigalovsky and Melcher, 2006), primates (Baumann et al., 2011; Kayser et al., 2007; Tanji et al., 2010) and songbirds (Boumans et al., 2007; Van Meir et al., 2005; Voss et al., 2007).

In this study, BOLD fMRI was applied to investigate the responses to UHF sounds in rat auditory system. Rat UHF vocalizations, single UHF pure tones, and pairs of UHF tones interleaved with pure tones equivalent to their quadratic distortion products were presented to the animals. The UHF values were systematically varied between 45 kHz and 75 kHz with 10 kHz steps, in order to examine the organization of UHF representation that may potentially exist. The findings were mainly focused on the IC and lateral lemniscus (LL), a subcortical auditory nucleus that projects to the IC (Malmierca, 2003).

Methods

Animal preparation

All animal experiments were approved by the local animal research ethics committee. Animals were prepared for fMRI experiments as described in our previous studies (Chan et al., 2010; Cheung et al., 2012a, 2012b; Gao et al., 2014; Lau et al., 2011, 2013, 2015; Zhang et al., 2013; Zhou et al., 2012, 2014). Briefly, male Sprague-Dawley rats (300–350 g, N = 10) were used in this study. They were initially anesthetized with 3% isoflurane and then mechanically ventilated via oral intubation. They were then placed on a holder in the prone position with a tooth bar to restrict head motion. Throughout the course of MR scanning, anesthesia was maintained with 1.2% isoflurane and warm water was circulated to maintain the animal temperature constant. Animal heart rate, respiration rate, oxygen saturation and rectal temperature were continuously monitored (SA Instruments).

Auditory stimulation

Auditory stimulation was controlled by a computer and produced by a pair of high-frequency magnetic speakers (MF1, TDT) driven by an amplifier (SA1, TDT) (Cheung et al., 2012a, 2012b; Gao et al., 2014; Lau et al., 2013, 2015; Zhang et al., 2013). Each speaker was connected to a 26 cm long rigid tube. The two tubes were merged by a Y-shape connector and catenated to a 144 cm long rigid tube and then a 6.5 cm long flexible tube. Monaural stimulation was delivered through this custom-made tube system (see illustration in Supplementary materials) into
the right ear of animals. The left ear was occluded with cotton wool and Vaseline, to reduce the scanner acoustic noise reaching the ears.

Animals were stimulated with a standard block design paradigm that consisted of 20 s sound-on and 40 s sound-off periods (Fig. 1). Nine fMRI sessions with different sound stimuli were performed on each animal, and each session was repeated twice. In Session 1, rat UHF vocalizations (obtained online from http://www.avisoft.com/rats.htm, the 2nd–5th of the “various 50 kHz rat calls”) were presented. To avoid confounding, background noise below 40 kHz in the original sound waveforms was filtered out, without affecting the spectra (between 45 and 70 kHz) of the vocalizations (Figs. 2A–D). Each of the four filtered waveforms was then played in a different block. Due to different lengths, they were repeated 160/320/200/200 times during the 20 s sound-on period, respectively, resulting in a duty cycle of 80%. In Sessions 2–5, single tones at 45 kHz, 55 kHz, 65 kHz and 75 kHz (Figs. 2E–H) were continuously presented, respectively. In Sessions 6–9, pairs of UHF pure tones (45 kHz & 55 kHz, 55 kHz & 65 kHz, 45 kHz & 65 kHz, 45 kHz & 75 kHz, Figs. 2I–L) and pure tones equivalent to their quadratic distortion products (i.e. 10 kHz, 10 kHz, 20 kHz and 30 kHz, Figs. 2M–O) were continuously presented in six interleaved blocks. Note that to present two pure tones simultaneously, both speakers were used, each playing one frequency. To present a single pure tone or a vocalization, only one speaker was used; the other one was inactivated to avoid cross-channel interactions.

The auditory stimulation was calibrated before each fMRI experiment outside the MRI magnet. The output waveform of each sound was measured by a recorder (FR2, Fostex, Japan) that was placed at ~2 mm from the tip of the flexible tube (Gao et al., 2014). The variance of the sound pressure level of each sound was maintained less than 2 dB and no distortion was generated by the sound delivery system.

MRI procedure

All MR experiments were performed on a 7T MRI scanner (PharmaScan 70/16, Bruker Biospin GmbH) using a transmit-only birdcage coil in combination with an actively decoupled receive-only surface coil. Scout images were first acquired to determine the coronal and sagittal planes of the brain. Ten coronal slices with 1.0/0.2 mm thickness/gap were then positioned to cover the major structures along the auditory pathway (Fig. 3A) with the third slice on the center of the IC (bregma – 8.5 mm, Fig. 3B). T2 weighted (T2W) images of these slices were acquired for anatomical reference using a Rapid Acquisition with Refocused Echoes (RARE) sequence (FOV = 32 × 32 mm², data matrix = 256 × 256, RARE factor = 8, TE = 36 ms, TR = 4200 ms, average = 2). FMRI measurements were obtained for the same slices using a single-shot Gradient-Echo Echo-Planar-Imaging (GE-EPI) sequence (FOV = 32 × 32 mm², data matrix = 64 × 64, flip angle = 56°, TE = 20 ms, TR = 1000 ms, NEX = 280 for Sessions 1–5 and 400 for Sessions 6–9).

FMRI data analysis

For each animal, all GE-EPI images were realigned to the mean image of the first MRI session using SPM8 (Wellcome Department of Imaging Neuroscience, University College London, UK). Then the T2W images from each animal were registered to a custom-made brain template acquired from a separate age-matched rat with the same settings. Registration was performed by affine transformation and Gaussian smoothing, with the criteria of maximizing normalized mutual information (SPM8). Afterwards, the fMRI images were resliced correspondingly. For individual animals, data from repeated fMRI sessions were averaged, in-plane smoothed (FWHM = one pixel) and high-pass filtered (128 s). Then a general linear model (GLM) was applied to calculate the response coefficient (β) maps for each stimulus. Student’s t-test was performed to identify activated voxels using the threshold p ≤ 0.05 (corrected for false discovery rate) and cluster size ≥ 4. Then at group level, images corresponding to the same fMRI session were averaged across all animals. Similar smoothing, filtering and GLM processes were applied to map responses. Note that for Sessions 6–9, the β map for each stimulus was further normalized by its maximum value, in order to compare the spatial patterns between the responses to each pair of UHFs and the pure tone equivalent to their quadratic distortion product.

Results

Fig. 3C displays the activation (β value) map of the vocalizations from a representative animal, overlaid on the T2W anatomical images. Rat brain templates that correspond to the center of each slice are also
presented to identify major auditory structures (Paxinos and Watson, 2005). The left IC and LL (contralateral to the side of stimulation) were activated robustly by the vocalizations. Fig. 3D presents the map computed from the averaged fMRI images of the ten animals. Activated structures include the IC, LL and the medial geniculate body (MGB). Specifically in the IC, voxels in both the dorsolateral and ventromedial area exhibited strong responses to the UHF vocalizations. Note that the results here were obtained by treating four vocalizations together as stimulation-on condition and contrasted to baseline condition. We also treated them as different stimuli to compute the responses to each individual vocalization. No obvious difference was observed among their responses (see Supplementary materials).

In contrast to the vocalizations, in Sessions 2–5, BOLD responses to single tones at 45 kHz, 55 kHz, 65 kHz and 75 kHz were not detected in any of the structures along the auditory pathway.

Fig. 4 compiles the results of Sessions 6–9 in a representative animal. The IC was robustly activated by all four pairs of UHF pure tones and the LL was activated by 45 kHz & 55 kHz. In the IC, a cluster of voxels (delineated with black boundary) in the ventromedial side exhibited the strongest responses for all four pairs tested, clearly showing combination sensitivity. Lateral to these voxels, the strongest secondary responses were observed to shift from the dorsolateral to the ventromedial side when the difference frequency of the two tones increased from 10 kHz to 30 kHz. Meanwhile, pure tones at 10 kHz, 20 kHz and 30 kHz activated both the IC and the LL. In the IC, when the frequency increased from 10 kHz to 30 kHz, voxels with the strongest responses shifted from the dorsolateral to the ventromedial side, consistent with the tonotopic organization of rat IC as demonstrated in previous studies (Cheung et al., 2012a, 2012b; Malmierca et al., 2008). It is important to note that the same cluster of voxels (delineated with blue boundary) showed the strongest secondary responses to each pair of UHF pure tones and the strongest responses to the tone at their quadratic distortion frequency.

Fig. 5 shows the results of Sessions 6–9 from the averaged data of the ten animals. The IC was activated by all four pairs of UHF pure tones and

Fig. 2. The spectrograms of sounds recorded at the tip of the flexible tube using a recorder (FR2, Fostex, Japan) outside the MRI magnet, including four different vocalizations (A, B, C, D), pure tones at 45 kHz (E), 55 kHz (F), 65 kHz (G) and 75 kHz (H), pairs of pure tones at 45 kHz & 55 kHz (I), 55 kHz & 65 kHz (J), 45 kHz & 65 kHz (K) and 45 kHz & 75 kHz (L), and pure tones at 10 kHz (M), 20 kHz (N) and 30 kHz (O). The color scale indicates the magnitude of the short time Fourier transform of the recorded signals. The sound pressure level (in dB) is indicated for each sound.
showed the same pattern as described in Fig. 4 for an individual animal. The LL was activated by three pairs of UHFs except 55 kHz & 65 kHz. The three pure tones at 10 kHz, 20 kHz and 30 kHz activated both the IC and LL and similar tonotopic organization was observed in the IC. Importantly, voxels that demonstrated the strongest secondary responses to each pair of UHF pure tones also exhibited the strongest responses to the tone at their quadratic distortion frequency (see more information about the BOLD temporal profiles in Supplementary materials). Each slice was overlaid with a rat brain template that corresponds to the center of the slice (Paxinos and Watson, 2005) to indicate major auditory structures and their subdivisions. LL: lateral lemniscus; MGM: medial geniculate medial; MGD: medial geniculate dorsal; MGV: medial geniculate ventral; bic: brachium of inferior colliculus; BIC: nucleus of brachium of inferior colliculus; ECIC: external cortex of the inferior colliculus; DCIC: dorsal cortex of the inferior colliculus; CNIC: central nucleus of the inferior colliculus; DNLL: dorsal nucleus of the lateral lemniscus; INLL: intermediate nucleus of the lateral lemniscus.

Discussion

The present fMRI results demonstrated that UHFs are represented in the auditory system with different mechanisms from tonotopy. The strong BOLD responses in the IC, LL and MGB to UHF vocalizations demonstrated that the auditory system is highly sensitive to UHFs. However, BOLD responses to single UHF tones, which would be naturally expected from this sensitivity, were not observed. Together with previous electrophysiological studies (Kelly et al., 1991; Malmierca et al., 2008), this indicated that few neurons are tuned to single UHF pure tones in the auditory system, i.e., UHFs are not integrated in tonotopy. On the other hand, robust BOLD responses in the IC and LL to pairs of UHF pure tones indicated that the auditory system encodes UHFs by combining information in different frequency components or exploiting the interactions between different frequencies. For four pairs of UHF pure tones tested systematically in this study, voxels in the ventromedial side of the IC always exhibited the strongest BOLD responses, indicating that a major or substantial population of neurons in these voxels is combination sensitive. Meanwhile, voxels in the dorsolateral side of the IC responded to each pair of UHFs in a similar way to a pure tone at their quadratic distortion frequency, suggesting that neurons in these voxels are activated by the distortion products. The results indicated that combination sensitivity and cochlear distortion are employed by large, but spatially distinctive neuron populations in the IC to represent UHFs.

UHF encoding in the IC

The present fMRI study demonstrated that combination sensitivity and cochlear distortion are adopted by different neuron populations in the auditory midbrain to encode UHFs. While both categories of neurons have been identified by previous electrophysiological studies,
BOLD fMRI revealed their different roles in auditory information encoding.

Combination sensitive neurons have been examined extensively in the IC of bats (Leroy and Wenstrup, 2000; Mittmann and Wenstrup, 1995; Nataraj and Wenstrup, 2005, 2006; Peterson et al., 2008; Portfors, 2004; Portfors and Wenstrup, 2002; Wenstrup et al., 2012) and mice (Felix and Portfors, 2007; Portfors and Felix, 2005). Such neurons exhibit either excitatory or inhibitory responses when two tones with distinctly different frequencies are presented simultaneously. It has been shown that facilitatory and inhibitory combination sensitivity originate in the IC and LL respectively (Wenstrup et al., 2012), with inhibition mechanisms and glycinergic inputs playing critical roles in creating such responses (Nataraj and Wenstrup, 2005, 2006; Wenstrup et al., 2012). Previous electrophysiological studies also reported similar percentage of facilitatory and inhibitory combination-sensitive neurons. In the current study, the positive BOLD responses to UHF tone pairs in the ventromedial IC probably resulted from the activation of facilitatory neurons. This is also supported by evidence in previous electrophysiological studies that combination sensitive facilitation is closely related to high frequency encoding. For example, in bat IC, facilitatory responses are commonly observed in representing frequencies above 30 kHz (Leroy and Wenstrup, 2000; Mittmann and Wenstrup, 1995; Nataraj and Wenstrup, 2005; Portfors and Wenstrup, 2002). In the IC of mice with high frequency hearing loss, less neurons display facilitatory combination sensitivity than in the IC of mice with normal hearing (Felix and Portfors, 2007; Portfors and Felix, 2005).

The bulk hemodynamic responses due to facilitatory combination sensitivity firstly indicated the existence of a large population of
combination sensitive neurons in the ventromedial side of the IC. In previous electrophysiological studies, the percentage of combination sensitive neurons in the IC varied substantially across studies and species (Leroy and Wenstrup, 2000; Mittmann and Wenstrup, 1995; Nataraj and Wenstrup, 2005; Peterson et al., 2008; Portfors and Felix, 2005; Portfors and Wenstrup, 2002). While this could arise from different testing and quantification methods adopted by different studies, it may also arise from the variation of sampling locations as suggested by the present results. In this study, it was unclear whether different ensembles of neurons were activated by different pairs of UHFs, and if so, how they were distributed within the combination-sensitive voxels. It was possible that some neurons responded to multiple pairs of UHFs as previous studies have found a small number of neurons that possess more than one facilitatory frequency band (Portfors and Wenstrup, 2002). Future investigations that can provide higher spatial resolution and sample a sufficient number of neurons may address this question. Secondly, the strong BOLD responses possibly suggested a primary role of combination sensitivity in encoding UHF information and processing species-specific vocalizations that contain them. In fact, it has been suggested that combination sensitivity is a fundamental mechanism for generating selective responses to acoustic signals at cellular level, and this underlies the functional specialization at cortical level that is key to the processing of human speech (Rauschecker and Scott, 2009).

During auditory stimulation with several frequency components, distortion products can be generated as a byproduct of nonlinear cochlear amplification (Ehret and Fischer, 1991). These products can

Fig. 5. Normalized activation (β value) map computed from the averaged fMRI images of the ten animals in Sessions 6 (A), 7 (B), 8 (C) and 9 (D). Activated voxels were identified by the threshold p < 0.05 (corrected for false discovery rate) and cluster size ≥ 4. The β map for each stimulus was normalized by its maximum value and those above 0.1 were color-coded. Similar to Fig. 4, two clusters of activated voxels in the left IC were delineated for each pair of UHF pure tones. Combination-sensitive voxels: a cluster in the ventromedial side (black boundary) that always exhibited the strongest responses to the four pairs of UHF pure tones tested. Distortion-sensitive voxels: voxels in the dorsolateral side (blue boundary) that exhibited strongest secondary responses to each pair of UHF pure tones and strongest responses to a pure tone at their quadratic distortion frequency, i.e., the difference between the two UHFs. The LL was activated by all stimuli except 55 kHz & 65 kHz. The activated voxels in the ventral part of the brain at 10, 20 and 30 kHz likely contained signals from multiple auditory nuclei, including the ventral nucleus of the LL, olivary complex and the trapezoid body. Note that only the slice covering the center of IC is shown here. BOLD responses in other slices were much weaker than this slice.
propagate along the cochlea to their CF site on the basilar membrane, and then excite neurons with the same CF at higher levels along the auditory pathway (Abel and Kossi, 2009; Faulstich and Kossi, 1999; Purcell et al., 2007). The present fMRI results suggested that cochlear distortion can be utilized by neurons in the IC that are tuned to lower frequencies to respond to simultaneously presented UHFs that individually cannot evoke responses. While similar phenomenon in single neurons has been demonstrated by recent electrophysiological studies (Portfors and Roberts, 2014; Portfors et al., 2009), the current imaging findings in a large spatial scale demonstrated that this mechanism is prevalent in the IC. Meanwhile, the BOLD responses evoked by cochlear distortion were relatively weaker than those that arose from combination sensitivity for each pair of UHFs, possibly suggesting that cochlear distortion plays a secondary role in UHF encoding in the IC.

It is likely that the involvement of multiple neural mechanisms enables the IC to encode UHFs and process complex sounds that contain them (e.g., vocalizations) more accurately (Holmstrom et al., 2010). It is plausible that the combination (i.e., summation) and distortion (i.e., difference) of two UHFs contain complementary signals, with which the IC can fully decode and extract information conveyed in both frequency channels. If only the combination or distortion were encoded, neurons might not accurately differentiate UHFs from lower frequencies. Such complexity may have been developed by animals while they evolved to emit UHF vocalizations to avoid predation, such that auditory neurons originally tuned to lower frequencies can encode their vocalizations with high efficiency and fidelity (Woolley and Portfors, 2013). The combination of different mechanisms also emphasized a critical role of the IC in processing UHF related sounds. This seems in line with the widespread anatomical connections between the IC and other auditory centers. By integrating all ascending signals from brainstem nuclei (Winer and Schreiner, 2005), the IC is in a pivotal position that can combine different inputs and perform fundamental processing before sending information to higher nuclei, i.e., the thalamus and cortex. It should be noted that other neural mechanisms besides the two discussed here may also be related to UHF encoding in the IC. For example, previous studies have identified multiply tuned neurons in bat IC with one excitatory band in the UHF range (Holmstrom et al., 2007; Portfors and Wenstrup, 2002). Such neurons likely constitute a small population, since BOLD responses to a single UHF pure tone were not detected in this study; yet their potential contribution could not be entirely ruled out. Further verification of the mechanisms demonstrated by the current results may be performed with electrophysiology in the future.

**BOLD responses to UHFs in the LL and other structures**

BOLD responses were detected in the LL for three pairs of UHFs examined in this study. For 55 kHz & 65 kHz, no significant responses were detected probably because the resulted neuronal activation was weak in nature. For the other three pairs, BOLD responses in the LL were generally homogeneous, without showing the pattern that characterizes activation induced by combination sensitivity and cochlear distortion as found in the IC. This may be explained by the missing of neurons that display combination sensitive facilitation in the LL (Wenstrup et al., 2012). Inhibitory combination sensitivity that originates in the LL (Wenstrup et al., 2012), by its nature, was not expected to evoke positive BOLD signals. It was also possible that the current imaging protocol was suboptimal for investigating LL responses. Future studies may examine LL encoding of UHF sounds in more details.

In other auditory nuclei, e.g. the MGB and the auditory cortex, BOLD activation was not observed for any of the four pairs of UHFs used in this study. This may also be explained by the weak nature of the responses in these structures given the low sound pressure level of the stimulation. In addition, the lack of BOLD responses in the auditory cortex might be related to the administration of anesthesia, which has been suggested to affect neural responses in the auditory cortex more than in the IC (Schumacher et al., 2011; Ter-Mikaelian et al., 2007). Notably, it has been documented that neurons in the auditory cortex of anesthetized animals display phasic responses to most acoustic stimulation, sometimes only a single spike per stimulus (Ter-Mikaelian et al., 2007; Wang et al., 2005). Sustained responses were observed rarely, e.g. when a preferred stimulus was presented (Wang et al., 2005). Since BOLD signals are highly correlated with neuronal firing rate (Logothetis et al., 2001), it was possible that phasic responses in most neurons would result in undetectable BOLD signal changes in the auditory cortex. In our previous auditory fMRI studies, BOLD responses in the MGB and auditory cortex were also lower than those in the IC and LL for other type of acoustic stimulation (Cheung et al., 2012a; Zhang et al., 2013). Future studies may investigate UHF encoding in these structures more carefully.

**BOLD responses to social vocalizations**

The present study presented large-view imaging results on responses to species-specific vocalizations in the rodent auditory system. This was in parallel with many electrophysiological studies that recorded responses to similar vocalizations at neuronal levels (Bauer et al., 2002; Mayko et al., 2012; Pincherli Castellanos et al., 2007; Suta et al., 2007; Wallace et al., 2005, 2013). The UHF vocalizations used in this study was in the category of “50 kHz calls” (the parallel category “20 kHz calls” was not examined in this study) (Brudzynski, 2009; Portfors, 2007). Although presented at relatively low sound pressure levels (below 60 dB, Figs. 2A–D), these sounds evoked robust BOLD responses in multiple auditory nuclei, including the IC, demonstrating that the auditory system is highly sensitive to these behaviorally relevant sounds (Romanski and Averbeck, 2009; Woolley and Portfors, 2013). Specifically in the IC, the strongest BOLD responses were observed in its dorsolateral region (Fig. 3). Note that these voxels likely contained neurons from both the external cortex of the IC (ECIC) and the CNIC (Paxinos and Watson, 2005). While the activation of CNIC neurons were not surprising given that cochlear distortion could map UHFs to lower frequencies (Portfors and Roberts, 2014; Portfors et al., 2009), the detailed mechanism for ECIC neurons to respond to these vocalizations remained to be further understood. This information further suggested that auditory encoding of complex vocalizations might engage additional circuits and neural mechanisms than those needed for processing two-tone stimulation. Furthermore, in the MGB, group-level BOLD responses were detected for the UHF vocalizations. Interestingly, the activated area included only the ventral part of MGB, but not the dorsal part (Fig. 3D). This may suggest that neurons in different MGB subdivisions exhibit different response properties to these vocalizations (Philibert et al., 2005; Suta et al., 2007), which might be further related to their different frequency tuning properties (Wenstrup, 1999).

Previous electrophysiological studies have discussed extensively the response selectivity of IC neurons among different vocalizations, i.e., the phenomenon that a specific neuron only responds to one or few vocalizations among many tested (Mayko et al., 2012; Pollak, 2013). In the current study, BOLD responses to the four vocalizations were similar (Supplementary materials), indicating that the bulk of neurons contained in each voxel have a low selectivity among these vocalizations. This may be caused by low response selectivity in single neurons, or a mixture of neurons that are selective to different vocalizations in each voxel. Interestingly, note that the vocalizations used in this study were recorded from a different rat strain (Wistar). The strong responses evoked in Sprague-Dawley rats may suggest that the auditory sensitivity to conspecific vocalizations does not require auditory experience (Mahrt et al., 2013).

**MEMRI study with two-tone stimulation**

In Yu et al., 2007, neonatal mice were exposed to a two-tone stimulus (16 kHz & 40 kHz) from postnatal days (P) 9 to 17 (Yu et al., 2007).
After this rearing period, manganese-enhanced MRI (MEMRI) was employed to map the accumulative IC activities (Yu et al., 2005) elicited by a subsequent 24 h pure tone stimulation (16, 32 or 40 kHz). Without two-tone stimulation during the manganese uptake stage, they did not report IC activities that might be evoked by cochlear distortion or arise from combination sensitivity. Note that compared to BOLD fMRI, which is based on the hemodynamic signals, MEMRI can provide more direct information of neural activities (Lin and Koretsky, 1997) and can be performed with potentially higher spatial resolution. Future studies therefore may combine BOLD fMRI and MEMRI to further characterize the spatiotemporal patterns of auditory brain responses to UHF sounds.

**Choices of frequencies in the two-tone paradigm**

In previous electrophysiological studies, when a neuron was tested for combination sensitivity or cochlear distortion using the two-tone paradigm, one frequency was usually fixed at its CF (Felix and Portfors, 2007; Mittmann and Wenstrup, 1995; Nataraj and Wenstrup, 2005, 2006; Portfors and Felix, 2005; Portfors and Wenstrup, 2002) or set in the frequency range of biosonar (Peterson et al., 2008) or a vocalization (Portfors et al., 2009) that the neuron was sensitive to; the other one was usually varied within the audible frequency range to search for the best facilitatory/inhibitory frequency or cochlear distortion products. This scheme was a reasonable and natural choice to reduce the workload given the infinite possibilities if both frequencies were to be set randomly. However, this may have limited their findings, because on one hand, other untested two-tone combinations were also possible to activate the neuron, and on the other hand, the frequency values covered in a study usually have a large variation. In this study, a different strategy was used. UHFs were chosen first and BOLD responses throughout the auditory system were then investigated. By systematically varying the frequency values, the current results enabled us to examine the mechanisms of UHF encoding in a different angle.

It is important to note that the mechanism of combination sensitivity may not only operate for two UHFs. Based on electrophysiological studies, they should also operate for combinations of an UHF and a lower frequency (Portfors and Felix, 2005; Portfors and Roberts, 2014; Portfors et al., 2009; Wenstrup et al., 2012). Such combinations were not examined in the present study, because a lower frequency itself would be expected to evoke BOLD responses, which would confound the combination- or distortion-evoked responses.

**Technical considerations**

When two tones with different frequencies are played, distortion generated by the stimulation setup must be prevented (Portfors and Felix, 2005; Portfors et al., 2009). In our study, measurement of the sound output from the sound delivery system confirmed that no distortion was generated before the stimulation entered the ear of the animals (Fig. 2).

During the sound-on period, the UHF tone pairs were played with 100% duty cycle. This was necessary because different frequencies take unequal time to propagate through the tubes. If small gaps were inserted during this period to reduce the duty cycle, potential confounding issues may arise from the asynchrony of the two tones. With 100% duty cycle, the two tones were simultaneously on throughout the 20 s period (except the asynchrony at initial onset and final offset). Furthermore, to prevent inter-channel crossovers, only a single speaker was used when a single sound was played.

In this study, the four UHF tones were played at different sound pressure levels (Fig. 2), which was mainly due to the limitation of the speakers used in this study. The speakers provided a large output range of ~80 kHz, but the peak output was lower for UHFs than lower frequencies (0–40 kHz: >90 dB; 40–80 kHz: gradually decreasing from 90 dB to 60 dB). In order to obtain robust BOLD responses for each pair of UHF tones, the highest intensity for each UHF tone that could be produced by the stimulation system was used (specified in Figs. 2E–H). The use of different intensities was not expected to influence the spatial patterns of the IC BOLD responses to each pair of UHF tones, as previous electrophysiological studies have not reported that the facilitatory frequencies of combination-sensitive IC neurons are dependent on sound intensities. Nevertheless, future studies may investigate the auditory responses to UHF sounds in terms of sound intensity in more details.

The current study did not employ sparse sampling paradigm (Hall et al., 1999; Talavage and Hall, 2012), as our recent work suggested that this is not a prerequisite for auditory fMRI studies (Cheung et al., 2012a, 2012b; Gao et al., 2014; Lau et al., 2013, 2015; Zhang et al., 2013). The continuous imaging protocol adopted here provided more time points for data analysis.

**Conclusion**

The present fMRI study demonstrated the robust sensitivity of the auditory system to UHF vocalizations, despite limited representation of single UHF pure tones in tonotopy. More importantly, with BOLD responses to pairs of UHF pure tones, the present study indicated that combination sensitivity and cochlear distortion are employed to represent UHFs by large, but spatially distinctive neuron populations, in the ventromedial and dorsolateral side of the IC, respectively. The sophisticated encoding mechanism in the IC highlighted a critical role of this subcortical structure in processing UHFs as well as complex sounds that contain them, such as vocalizations. These imaging findings also provided insights for understanding sound feature encoding in the early stage of the auditory pathway.

**Abbreviations**

UHF ultrahigh frequency  
CF characteristic frequency  
IC inferior colliculus  
LL lateral lemniscus  
MGB medial geniculate body  
BOLD blood-oxygen-level-dependent  
fMRI functional MRI  
T2W T2 weighted  
GLM general linear model

**Acknowledgments**

The authors would like to thank Professor Dan H. Sanes from New York University for his insightful comments on the manuscript. This work was supported in part by Hong Kong Research Grants Council (HKU783/11M), Croucher Foundation for Senior Researcher Fellowship (EXW) and Lam Woo Foundation for Endowed Professorship (EXW).

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.04.007.

**References**


Wenstrup, J.J., Nataraj, K., Sanchez, J.T., 2012. Mechanisms of spectral and temporal integra-

tion in the mustached bat inferior colliculus. Front. Neural Circ. 6, 75.


New York, NY.

Woolley, S.M., Portfors, C.V., 2013. Conserved mechanisms of vocalization coding in mam-


mice with Mn-enhanced MRI. Nat. Neurosci. 8, 961–968.

nization of the tonotopic map in mouse auditory midbrain revealed by MRI. Proc.


the auditory system in rats. Proceedings of the 17th Annual Meeting of ISMRM,

Honolulu, USA, p. 1605.


Functional magnetic resonance imaging of sound pressure level encoding in the rat


68, 65–73.


resting-state functional MRI connectivity: morphological foundation and plasticity.

Neuroimage 84, 1–10.