J Physiol 599.10 (2021) pp 2527–2557

The Journal of Physiology

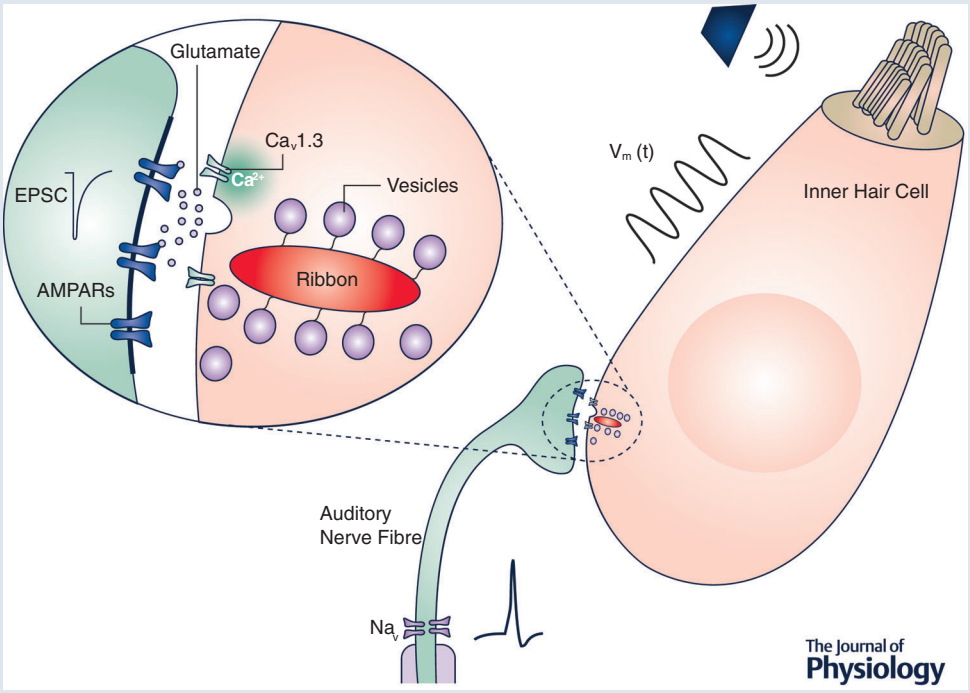
TOPICAL REVIEW

Encoding sound in the cochlea: from receptor potential to afferent discharge

Mark A. Rutherford[1], Henrique von Gersdorff[2] and Juan D. Goutman[3]

[1]*Department of Otolaryngology, Washington University School of Medicine, St Louis, MO, 63110, USA*
[2]*Vollum Institute, Oregon Hearing Research Center, Oregon Health and Sciences University, Portland, OR, 97239, USA*
[3]*INGEBI (CONICET), (1428) C. A. Buenos Aires, Argentina*

Edited by: Ian Forsythe & Walter Marcotti



Abstract Ribbon-class synapses in the ear achieve analog to digital transformation of a continuously graded membrane potential to all-or-none spikes. In mammals, several auditory nerve fibres (ANFs) carry information from each inner hair cell (IHC) to the brain in parallel. Heterogeneity of transmission among synapses contributes to the diversity of ANF

Mark A. Rutherford is an Assistant Professor in the Department of Otolaryngology at Washington University in St Louis, MO, USA. Previous appointments include postdoctoral fellowships with Tobias Moser in the Inner Ear Lab at the University of Göttingen, Germany, and with William Roberts at the University of Oregon in Eugene. **Juan D. Goutman** is an Independent Researcher in the National Scientific and Technical Research Council at the Universidad of Buenos Aires, Argentina. Previous appointments include a postdoctoral fellowship with Elizabeth Glowatzki and Paul Fuchs at Johns Hopkins Medical School in Baltimore, Maryland. Like their co-author, **Henrique von Gersdorff**, they have dedicated their careers in research to understanding the biophysical mechanisms of ribbon synapses.

DOI: 10.1113/JP279189

sound-response properties. In addition to the place code for sound frequency and the rate code for sound level, there is also a temporal code. In series with cochlear amplification and frequency tuning, neural representation of temporal cues over a broad range of sound levels enables auditory comprehension in noisy multi-speaker settings. The IHC membrane time constant introduces a low-pass filter that attenuates fluctuations of the receptor potential above 1–2 kHz. The ANF spike generator adds a high-pass filter via its depolarization-rate threshold that rejects slow changes in the postsynaptic potential and its phasic response property that ensures one spike per depolarization. Synaptic transmission involves several stochastic subcellular processes between IHC depolarization and ANF spike generation, introducing delay and jitter that limits the speed and precision of spike timing. ANFs spike at a preferred phase of periodic sounds in a process called phase-locking that is limited to frequencies below a few kilohertz by both the IHC receptor potential and the jitter in synaptic transmission. During phase-locking to periodic sounds of increasing intensity, faster and facilitated activation of synaptic transmission and spike generation may be offset by presynaptic depletion of synaptic vesicles, resulting in relatively small changes in response phase. Here we review encoding of spike-timing at cochlear ribbon synapses.

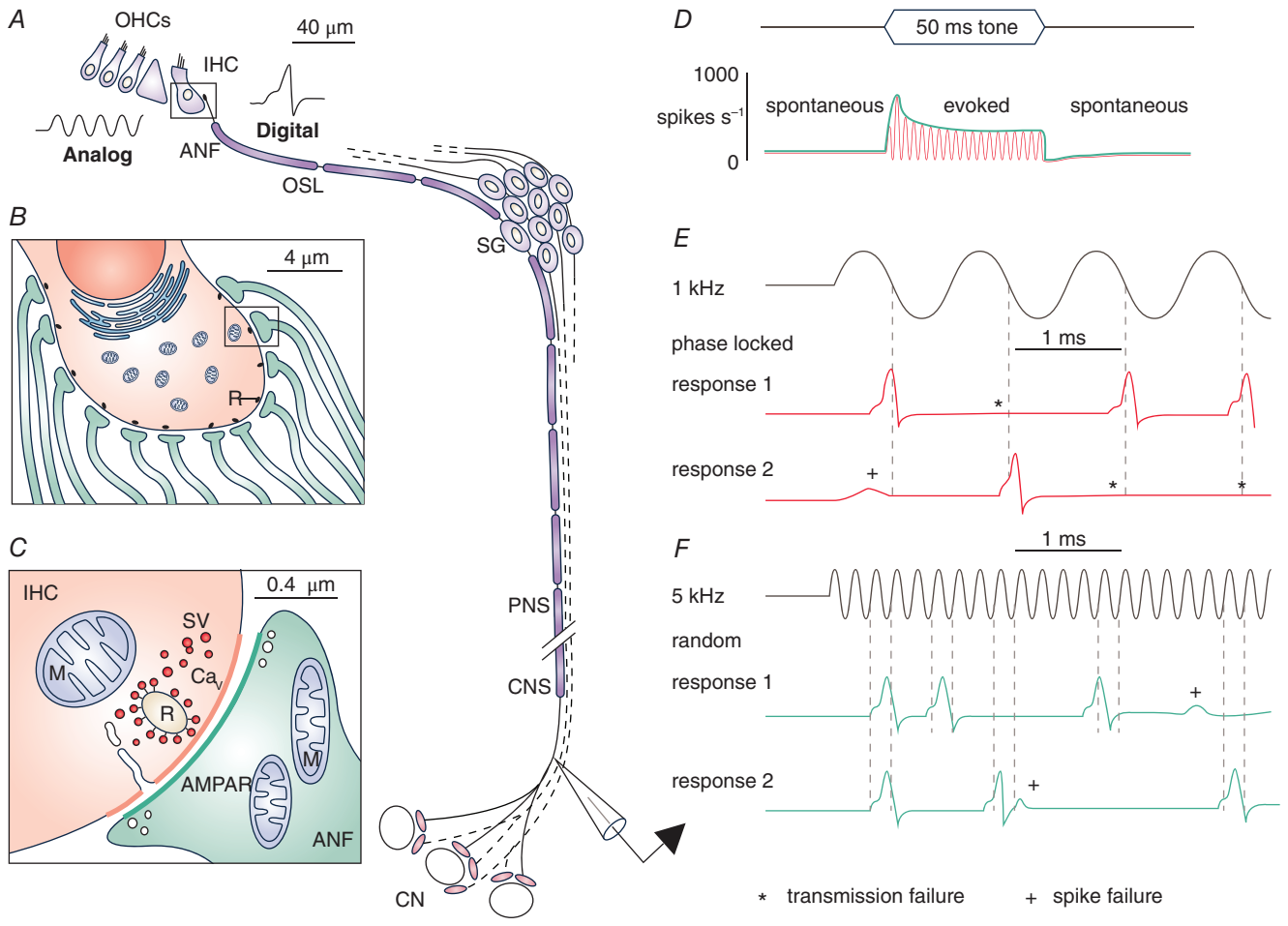
(Received 22 July 2020; accepted after revision 22 February 2021; first published online 28 February 2021)
Corresponding author Mark A. Rutherford: Department of Otolaryngology, Box 8115, Washington University in St Louis School of Medicine, 660 S. Euclid Ave., St Louis, MO 63110, USA. Email: RutherfordMark@wustl.edu

Abstract figure legend The auditory ribbon synapse controls the timing of afferent spikes. The afferent synapses between sensory hair cells and primary auditory neurons are glutamatergic ribbon-class synapses. This schematic representation highlights a ribbon synapse between a cochlear inner hair cell and one of its auditory nerve fibres (ANF). Sound results in mechanical stimulation of the inner hair cell bundle, evoking a transduction current that generates a receptor potential $V_m(t)$ that changes over time with sound level and frequency. Depolarization of $V_m(t)$ evokes Ca^{2+}-dependent exocytosis of glutamate at the ribbon synapse. Glutamate binds to AMPA-type receptors to generate an excitatory postsynaptic current (EPSC) in the ANF dendrite. The ensuing depolarization triggers an action potential in the ANF, signalling sound information to the brain.

Abbreviations AM, amplitude modulation – change in envelope of sound pressure waveform over time; ANF, auditory nerve fiber, type-I – primary afferent neuron, excited by cochlear ribbon synapse, also called type-I spiral ganglion neuron; AZ, active zone – presynaptic structures mediating neurotransmitter release; CF, characteristic frequency – tone frequency of greatest sensitivity for a given ANF; C_m, membrane capacitance – electrophysiological property of the cell's plasma membrane, proportional to surface area; EPSC, excitatory postsynaptic current – measurement of neurotransmission from IHC to ANF; EPSP, excitatory postsynaptic potential – depolarization of ANF caused by the EPSC; f_c, corner frequency – above which the response is reduced to ∼70% of maximum; I_{Ca}, calcium current – presynaptic voltage-gated transmembrane current; I_{Tr}, transduction current – mechanically gated at the tips of hair cell stereocilia, see MET; IHC, inner hair cell – primary sensory receptor of the cochlea, excited mechanically by fluid displacement of stereocilia, presynaptic to several ANFs; OHC, outer hair cell – mechanically sensitive amplifier of sound-evoked vibrations, responsible for cochlear sensitivity and sharpness of tuning; MET, mechano-electric transduction – channel activated by vibration of stereocilia in response to sound, mediator of I_{Tr}, partly open at rest; P_r, probability of release – of a glutamate-filled synaptic vesicle from IHC to ANF; R_m, membrane resistance – electrophysiological property of the cell's plasma membrane; SGN, spiral ganglion neuron (ANF), type-I – bipolar neuron connecting organ of Corti to brain, comprising 95% of the afferent fibers in the auditory nerve; SPL, sound pressure level – wave strength, correlates with perception of loudness; $P_s(t)$, spike rate as a function of time – instantaneous spike rate or mean rate over time; τ_m, membrane time constant – property of the cell membrane, R_mC_m; $V_m(t)$, membrane voltage as a function of time – hair cell receptor potential; $W(t)$, time varying amplitude of sound pressure waveform at the eardrum

Introduction

From ear to brain. The initial neural representation of sound is generated in the inner ear and conveyed to the brain as sequences of action potentials (spikes) across an array of primary auditory neurons in the cochlea called type-I spiral ganglion neurons (SGNs) or auditory nerve fibres (ANFs), comprising 95% of the afferent fibres in the auditory nerve (Spoendlin, 1972). Each ANF has a myelinated peripheral axon with a spike-generating heminode excited by a single afferent synapse on an inner hair cell (IHC). Cochlear afferent synapses (i.e. IHC–ANF synapses) contain a presynaptic cytoplasmic density called the synaptic ribbon, which harbours a pool of glutamate-filled synaptic vesicles. As shown in the Abstract Figure and Fig. 1,



The Journal of Physiology

Figure 1. Sound is encoded as spikes at the cochlear hair cell afferent synapse
A, in the organ of Corti, inner hair cells (IHCs) are stimulated by the movement of cochlear fluids, amplified and tuned by sound-evoked activity of the outer hair cells (OHCs). The graded transmembrane potential of the IHC (analog) is encoded as spikes (digital), generated at the nearby heminode on the auditory nerve fibre (ANF). The ANF peripheral axon is myelinated. Spikes are propagated at nodes of Ranvier in the osseous spiral lamina (OSL), across the soma in the spiral ganglion (SG), and along the central axon towards the brain. In the cochlear nucleus (CN), the ANF central axon branches to terminate on several principal cells. Extracellular sharp-electrode recordings of spike trains in ANFs are generally made from axons as they enter the brainstem. Abbreviations: CNS, central nervous system; PNS, peripheral nervous system. *B*, expansion of boxed area in *A*, showing the basolateral area of one IHC with its afferent ribbon synapses. Each ANF receives excitatory glutamatergic input from a single ribbon synapse (R, ribbon). The ANF dendrites are unmyelinated within the hair cell epithelium where they also receive efferent synaptic input (not shown). *C*, expansion of boxed area in *B*, showing one afferent synapse containing the presynaptic ribbon and the voltage gated Ca^{2+} channels in the membrane (Ca_V, thick orange line) that control release of synaptic vesicles (SV). Postsynaptic AMPA-type glutamate receptors in the membrane (AMPARs, thick green line) are required for the synaptic transmission that mediates hearing. M, mitochondria. *D*, upper, a silent period is interrupted by a 50 ms tone burst at 1 kHz or 5 kHz. Lower, idealized post-stimulus time histograms; representations of average responses of a single ANF to repetitions of a 1 kHz tone (red) or a 5 kHz tone (green), constructed by recording responses to *N* repetitions of the same 50 ms tone. The spike sequences are aligned at stimulus onset, the observation period is divided into time bins of size *T*, and spikes are counted in each bin. The mean spike rate per second for each bin is calculated by dividing the number of spikes per bin by $N \times T$. Before and after the tone burst, the spontaneous spike rate is observed. At sound onset, the spike rate increases to a peak rate that rapidly adapts to a steady-state rate primarily due to depletion of synaptic vesicles. The response to the 1 kHz tone is synchronized to the periodicity of the tone (red); the response to the 5 kHz tone is not (green). *E*, upper, approximately 3.5 cycles of a 1 kHz tone. Lower, ANF spike trains for two presentations of the 1 kHz tone demonstrate responses at a preferred phase (i.e. phase-locking). The ANF does not fire on every stimulus cycle

due to failures of synaptic transmission (∗) or failures of spike generation (+). The vertical dashed lines refer to the same arbitrary phase within each cycle. Imperfect synchronization results from the temporal jitter in synaptic transmission. *F*, upper, approximately 18 cycles of a 5 kHz tone. Lower, ANF spike trains for two presentations of the 5 kHz tone demonstrate apparently random timing of individual spikes relative to the phase of the stimulus. Synaptic transmission fails to evoke a spike (+) if the excitatory postsynaptic potential (EPSP) is too slow or too small (spike train 1) or if the EPSP follows a spike at very short latency while the ANF is refractory (spike train 2). Failures of synaptic transmission occur on every cycle lacking an EPSP or spike, but asterisks are omitted for clarity. The vertical dashed lines refer to the same arbitrary phase for two adjacent periods around each spike. Lack of synchronization results from low-pass filtering of the IHC membrane time constant and from temporal jitter in synaptic transmission.

each ribbon synapse is separated from its spike-generating heminode by only ∼20 μm of unmyelinated, unbranched ANF dendrite (Liberman, 1980). Mature ANFs do not spike endogenously, but ribbon synapses drive Ca^{2+}-dependent glutamate release in silence, evoking so-called 'spontaneous' spike trains in the postsynaptic neurons. When a synaptic vesicle is exocytosed from an IHC, glutamate binds to postsynaptic α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-type glutamate receptors (AMPARs) on the ANF postsynaptic terminal. Opening of AMPARs produces an excitatory postsynaptic current (EPSC). The ensuing excitatory postsynaptic potential (EPSP) depolarizes the nearby heminode, where a high concentration of voltage-gated Na^+ channels (Na_V) reside, where the myelin begins and where spikes are initiated. Spikes initiated distally near the ribbon synapse are propagated along the myelinated peripheral axon, through the SGN cell body in the spiral ganglion, and along the myelinated centrally projecting axon to the brainstem (Fig. 1*A*). Here we review aspects of signal processing in the mechanisms of receptor potential generation, synaptic transmission and spike generation in sensory hair cells and the auditory nerve. We focus on the temporal encoding of sound in IHCs and ANFs of the mammalian cochlea, from receptor potential to afferent discharge, making some comparisons with other vertebrates. Cochlear mechanics, mechano-transduction and temporal encoding in the auditory brainstem are amply reviewed elsewhere (Joris *et al*. 2004; Hudspeth, 2014; Reichenbach & Hudspeth, 2014; Fettiplace, 2017). For development of cochlear anatomy and function, ANF diversity, the efferent system, and genetics of hearing and deafness, the reader is referred to additional reviews (Guinan, 2010; Richardson *et al*. 2011; Bulankina & Moser, 2012; Johnson *et al*. 2019; Petitpré *et al*. 2020).

Hair cells detect sound-evoked vibrations of the basilar membrane. Activity at the ribbon synapse is modulated by the frequency, level and timing of sound. Each IHC responds to a range of sound frequencies and is most excitable at a characteristic frequency (CF) determined by mechanical filtering at its 'tonotopic' (i.e. frequency–place) position along the basilar membrane (Von Békésy & Wever, 1960; Greenwood, 1990; Müller, 1996; Tsuji & Liberman, 1997; Müller *et al*. 2005). ANFs innervating IHCs in the cochlear apex respond best to low frequencies while those from the base of the cochlea respond best to high frequencies (CF ranges approximately 20 Hz–20 kHz for humans). CF defines a position on the basilar membrane along the cochlea's tonotopic axis. ANFs with identical frequency tuning can differ from each other in several properties including spontaneous rate (SR), excitability threshold (the minimum sound pressure level (SPL) required to evoke an increase in spike rate), and dynamic range sensitivity (the range of SPL over which the ANF spike rate changes). Some of these properties show co-variations: high-SR ANFs tend to respond to the softest sounds (low-threshold) and have narrower dynamic ranges, whereas low-SR ANFs tend to respond only to louder sounds (high-threshold) and have broader dynamic ranges (Liberman, 1982). Thus, each ANF in the cochlear array is a labelled line sensitive to a sub-range of sound frequencies and levels. At the same time, information about sound level and frequency is contained in the timing of spikes in each of these labelled lines. The question we are concerned with here is how cellular mechanisms in the cochlea contribute to spike timing in ANFs. What mechanisms enable and limit the encoding of temporal information in the cochlea?

Mature mammalian hair cells, and photoreceptors in the retina that also contain ribbon synapses, do not fire action potentials but rather undergo graded changes in their transmembrane 'receptor potentials'. The graded receptor potential modulates Ca^{2+} influx and vesicle release at the ribbon synapses. The auditory information contained in the receptor potential of a single IHC diverges into the spike trains of 10–30 ANFs in an analog to digital transformation at the ribbon synapses (Fig. 1*A*–*C*). Diversity of sound-response properties among ANFs is essential for the full representation of sound in the auditory system (Kiang *et al*. 1965; Kiang, 1984). Heterogeneities among ribbon synapses within IHCs seem to be essential to this diversity (Meyer & Moser, 2010; Wichmann & Moser, 2015). In addition, neuron-intrinsic heterogeneity in ion channel expression likely contributes to diversity of ANF response properties (Petitpré *et al*. 2018; Shrestha *et al*. 2018; Sun *et al*. 2018). As well, dendrites of ANFs are targeted by modulatory neurotransmission from efferent neurons that can regulate

conductances and firing properties of ANFs (Nouvian *et al*. 2015).

In this topical review, only occasional reference will be made to the extensive research on timing and binaural hearing at the next stages of signal processing in the ascending auditory pathway. The auditory signals from IHC–ANF synapses are transmitted via the auditory nerve to the cochlear nucleus in the brainstem, where each ANF branches to form multiple presynaptic terminals (Fig. 1*A*). These glutamatergic projections are organized anatomically and functionally into distinct sub-nuclei onto at least six different types of principal cell (Brawer *et al*. 1974; Muniak *et al*. 2016). For example, in circuits specialized for temporal encoding in the ipsilateral anteroventral cochlear nucleus (AVCN), ANFs terminate onto somata of bushy cells via large axo-somatic synapses (endbulbs of Held) containing several hundred release sites per terminal (Xu-Friedman & Regehr, 2004; Moser and Strenzke, 2015; Wichmann, 2015; Rubio, 2018). Therefore, in each mammalian ANF, where spike timing is determined by activity at a single ribbon synapse, the timing of each spike may determine the timing of transmitter release at thousands of release sites or active zones (AZs) in the cochlear nucleus. Moreover, the ANF spike can be triggered by release of a single synaptic vesicle, exocytosis of which may be determined by the timing of opening of just one or very few Ca^{2+} channels (Brandt *et al*. 2005; Wong *et al*. 2014; Magistretti *et al*. 2015; Rutherford and Moser, 2016). In this way, the opening of a single Ca^{2+} channel in the ear may trigger exocytosis of thousands of vesicles at the next stage of sensory processing in the cochlear nucleus. How is the information in sound encoded temporally and what is the relevance of temporal encoding to perception?

Peripheral spectral and temporal processing. Spectral information is mainly encoded in the spatial distribution of ANF activity, as the cochlea performs a Fourier analysis on complex sounds to signal the relative amplitudes of frequency components to the brain. In contrast, temporal information is carried through precise timing of spikes within and among groups of fibres. The importance of spectral *vs*. temporal processing has been long debated. With spectral processing, sound frequency is encoded by the cochlear place of excitation. In temporal processing, frequency is encoded by response periodicity. Both seem to be important, as the auditory system appears to employ spectral–temporal processing, consisting of tonotopic representation combined with temporal analysis to extract information about the spectrum of acoustic signals over time (Eggermont, 2001). The importance of spectral processing to auditory encoding is reflected in the existence of tonotopic organization at all stages of the auditory pathway from the periphery to the cortex (Suga, 1965; Shamma, 2001). The human cochlea contains about 3000 IHCs, representing as many frequency channels to encode the spectrum of sound. In contrast, modern cochlear implants contain about 10 stimulating electrodes. Although speech intelligibility depends on spectral and temporal features of sound, with spectral cues increasing in importance in difficult listening environments, the significance of temporal processing is evidenced by the ability of cochlear implants to enable speech perception despite relatively poor spectral resolution (Shannon, 2002; Nie *et al*. 2006; Oxenham & Kreft, 2014).

The mammalian cochlea evolved for processing a wide range of complex sounds, some with pseudo-periodic components from tens of hertz to a few thousand hertz. Complex sounds are broken down into frequency bands that convey information over time scales ranging from hundreds of milliseconds to <1 ms. Temporal information in speech can be classified into three categories: fine structure – variations in wave shape over short intervals, within a single period of periodic sounds, with fluctuation rates from 500 to 10,000 cycles s^{-1}; periodicity – usually due to the carrier or fundamental frequency, with fluctuation rates from 20 to 500 cycles s^{-1}; and temporal envelope – changes in overall amplitude of the sound waveform, with fluctuations generally from 2 to 20 cycles s^{-1}. Temporal envelope and temporal fine-structure cues underlie speech identification, with fine-structure cues becoming more important in difficult listening environments (Lorenzi & Moore, 2007). Although generally existing in different time scales, these categories of temporal information are not strictly mutually exclusive, as temporal envelopes may be periodic. Like spectral information and tonotopic representation, the importance of periodic temporal envelope information to auditory encoding is reflected in the existence of periodotopic organization at all stages of the auditory pathway from the periphery to the cortex (Langner *et al*. 1997).

The temporal envelope of sound is like the output of a low-pass filter operation on the sound waveform. Periodic envelope fluctuations result when the frequency components of sound combine in the time domain to produce amplitude modulation (AM) of the sound waveform. For example, a single sinusoidal waveform can be amplitude modulated by a lower frequency sinusoidal envelope. For vowel production, cyclic opening and closing of the glottis at a carrier frequency produces pressure pulses that excite resonant modes of the vocal tract, producing harmonically structured spectral peaks, called formants. These spectral components combine in the time domain to produce phonemes, the components of words, which are pseudo-periodic envelope components with AM frequencies in the range of a few hertz to several hundred hertz. AM of the sound waveform is converted to AM of basilar membrane motion in cochlear regions

tuned to the spectral peaks. Although the AM components are not present as acoustic energy in the frequency spectrum of the sound pressure waveform, ANF spike timing may encode the frequency of AM if the spectral components fall within the frequency tuning curve of the ANF. Moreover, the frequency of AM may be perceived as a pitch, even though the frequency corresponding to that pitch is not present in the acoustic energy spectrum, as in the phenomenon of the missing fundamental frequency (Schouten, 1970; McFadden, 1988).

Spikes in ANFs may synchronize to periodic components of sound's temporal envelope or its fine structure up to a limit around a few kilohertz, but synchrony is stronger at lower frequencies (Palmer, 1982; Langner, 1992; Joris *et al*. 2004). With increasing frequency, spike-timing becomes random with respect to the periodicity of the stimulus (Fig. 1*D*–*F*). One reason for this upper frequency limit is bioelectrical impedance of the hair cell membrane, as discussed below in the section 'Membrane time constant'. Additional reasons, mainly due to the kinetics and jitter of synaptic transmission, are discussed in the section 'EPSC and spike latency, time course, and jitter'. Interestingly, the frequency limit for synchronization to AM is about one octave lower than the frequency limit for synchronization to pure tones (Joris & Yin, 1992; Greenwood and Joris, 1996). This may result from differences in the way hair cells are stimulated by pure tones *vs*. by AM. For low-frequency tones, the basilar membrane moves up and down and the hair bundle is positively and negatively displaced on alternating half-cycles of rarefaction and compression at the frequency of the tone. Relative to its resting potential, the hair cell is depolarized by the positive deflections and hyperpolarized by the negative deflections, which increases and decreases the probability of vesicle release relative to the spontaneous rate on alternating half-cycles of the oscillation. For a high-frequency tone (carrier frequency) sinusoidally amplitude modulated at low-frequency, there is no compression and rarefaction at the frequency of AM. Instead, the high-frequency bundle oscillation is modulated from peak-amplitude to zero and then back to peak-amplitude at the AM frequency. As explained below, due to the impedance of its plasma membrane and the asymmetry of the transduction current the IHC may not hyperpolarize from its resting potential during the AM stimulus. Half-cycle reduction in depolarization in the case of AM, rather than half-cycle hyperpolarization in the case of the tone, may result in less phase preference for the case of AM.

Temporal code and rate code. Spike patterns contain a rate code and a temporal code. Whereas a rate code is averaged over time, a temporal code is concerned with precise spike timing or rapid fluctuations in firing rate. These two concepts merge when a rate code is averaged over small times, such as a tone half-cycle. Although some aspects of temporal encoding are well described, for example phase-locking to tones, it is unclear how the brain uses this information to produce perception. As discussed in the section 'Spike timing precision in the ascending auditory pathway', behavioural discrimination thresholds are far more precise than the temporal precision of spike times observed in individual neurons. In analogy to the visual system, although we understand the biochemical mechanism governing signal transduction in the retina, it is unclear how signals from only three types of cones (red, green, blue) underlie discrimination of hundreds of colours.

Strictly periodic sounds like tones, while not found in nature, are useful for characterizing response functions. For tones of low frequency, in which several spikes may occur within one period, sound level may be encoded by the mean spike rate in an individual ANF (in addition to the number of active ANFs and their identities as high or low threshold). At the same time, the distributions of spike intervals could encode sound frequency in a temporal code, giving rise to the perception of periodicity pitch (Licklider, 1954). For AM sounds of low frequency, the same principles apply for extracting sound level from spike rates and sound frequency from spike intervals. ANFs have maximum spike rates of several hundred spikes per second. For tones or temporal envelope fluctuations greater than a few hundred hertz, a rate code averaged over several cycles could encode level.

Within frequency limits, below 1–3 kHz for most species, the temporal envelope of sound and the fine-structure of sound is represented in the temporal pattern of spikes in each ANF (Rose *et al*. 1967). For a given ANF, synchronization to temporal fine-structure or envelope depends on tonotopic position/CF and the relative levels of the signal components (Delgutte, 1980). Temporal fine-structure can be defined as the fast changes in the sound waveform utilized to perceive pitch, localize sounds and separate sound sources binaurally. Pitch may be encoded spectrally by tonotopic cues and temporally by periodicity cues. In cochlear implant users, pitch perception is poor due to (1) imprecise tonotopic resolution – individual electrodes stimulate broad frequency regions, and (2) limited temporal resolution – conventional signal processors filter out the temporal fine-structure, and the encoding of periodicity pitch is limited to several hundred hertz (middle C on a piano is 260 Hz). The highest key on a normal piano is ∼4 kHz, and a standard audiogram goes up to 8 kHz. In normal-hearing listeners, pitch perception is less robust above a few kilohertz due to (1) loss of temporal information when fluctuations are faster than the speed limits of ANF spike synchronization, and (2) the logarithmically spaced frequency–place map resulting

in less real estate for representation of higher frequencies. However, the auditory system can still make use of high-frequency temporal fine-structure up to ∼20 kHz as harmonics combine in the time domain to produce AM at a fundamental frequency.

Thus, ANFs are labelled lines that respond over ranges of sound frequency and level, and the pattern of spikes contains both a rate code for slow changes and a temporal code for faster changes in the sound pressure waveform. The precise spike train in each ANF is determined by the properties of the sound as well as the spectral- and level-dependent filtering unique to each ANF (Heil & Peterson, 2015). Temporal correlations of activity between groups of ANFs in spectrally resolved frequency bands contain information that could be used by the auditory system to detect signals in background noise or to distinguish two simultaneous speakers (Kim & Molnar, 1979; Carney, 2018; Huet *et al.* 2018). Although significant ANF loss can occur without affecting threshold for detection of low-level sounds, precise behavioural discrimination of sound frequency and level utilizes the full complement of IHCs and ANFs. We will now consider the stages of signal processing from hair cell receptor potential to neuronal spike generation.

Hair cell mechanotransduction and receptor potential

Mechano-electric transduction current. Positive displacements of the hair bundle evoke an inward mechano-electric transduction (MET) current that is graded with the size of the displacement (Fig. 2*A* and *B*). A fraction of the channels are open at rest, and negative deflections close those channels. The number of MET channels open at rest determines the size of the resting or 'silent' current which depolarizes the hair cell away from the K^+ equilibrium potential of ∼−80 mV towards that of the non-specific cationic MET conductance of ∼0 mV (Zidanic & Brownell, 1990; Johnson *et al.* 2011). The size of the silent current is determined by the open probability of the MET channels, which depends strongly on the ionic composition of the fluid bathing the hair bundle (Farris *et al.* 2006). In concert with basolateral K^+ conductance, the silent current determines the hair cell resting potential and the resting voltage-gated conductance of the cell, which has a major influence on hair cell receptor potential and overall cochlear responses to sound by determining the membrane time constant and the symmetry of evoked MET currents (Fettiplace, 2017). Hair cells contact two different fluids. The hair bundle is bathed in endolymph which has an unusual composition (low Ca^{2+}, high K^+) to support transduction. The basolateral membrane is bathed in perilymph, which is similar to cerebrospinal fluid (low K^+, high Na^+, high Ca^{2+}), to support Ca^{2+}-dependent exocytosis from the IHC and Na^+-dependent spike generation in the ANF. Thus, K^+ has a depolarizing influence as it enters through the MET channels and a hyperpolarizing influence as it exits through basolateral voltage-gated channels. High Ca^{2+} (∼1.3 mM in perilymph) blocks and lowers the open probability of the MET channels and reduces the size of the silent current, resulting in hyperpolarization and asymmetry of the current–displacement function around the resting position of the bundle (Fig. 2*B*, continuous curve; Corey & Hudspeth, 1983; Ricci *et al.* 2003, 2005). Low Ca^{2+} (∼20 μM in endolymph) increases MET open probability to ∼0.5, resulting in depolarization and a symmetric current–displacement function (Fig. 2*B*, dotted curve; Beurg *et al.* 2010; Johnson *et al.* 2011). MET currents are also shaped by adaptation of the current–displacement function, whereby increased intracellular Ca^{2+} may reduce the open probability (Ricci *et al.* 1998; Wu *et al.* 1999). In this way, intracellular Ca^{2+} buffering is thought to influence the MET silent current and thus the symmetry of evoked MET currents, as well as the resting potential and thus the membrane conductance and time constant (Johnson *et al.* 2011).

Membrane time constant. Changes in the sound pressure waveform at the eardrum over time ($W(t)$) are converted to changes in the hair cell receptor potential over time ($V_m(t)$) as vibrations of the middle ear and then cochlear fluids mechanically gate transduction currents ($I_{Tr}(t)$) through the hair bundle. However, gradations of $V_m(t)$ are limited in rate (mV/ms) by the low-pass filter property of the membrane resistance (R_m) and capacitance (C_m). This defines the membrane RC time constant ($\tau_m = R_m \times C_m$) (Fig. 2*A* and *C*). Changes in $W(t)$ slower than the corner frequency f_c ($f_c = 1/2\pi\tau_m$) modulate $V_m(t)$ on a cycle by cycle basis. Changes in $W(t)$ faster than the corner frequency are attenuated in $V_m(t)$ because of the time required to charge the electrical capacitance of the cell membrane, C_m. Membrane capacitance is proportional to surface area, and thus smaller cells change transmembrane voltage, V_m, faster. Membrane resistance, R_m, is inversely proportional to membrane conductance, and thus cells with greater conductance change V_m faster. The membrane conductance is determined primarily by voltage-gated K^+ channels in the basolateral membrane and the MET channels in the stereocilia. The K^+ conductances are carried predominantly by KCNQ4 ($K_V7.4$) and KCNMA1 channels ($K_{Ca}1.1$), called $I_{K,n}$ and $I_{K,f}$ in hair cells, as well as several other channels comprising $I_{K,s}$ (Kros & Crawford, 1990; Marcotti and Kros, 1999; Dierich *et al.* 2020). All these K^+ channels are thought to be at least partly open at the resting potential, and, as discussed above, hair cells with greater silent current through MET channels will have more depolarized resting potentials and greater membrane

conductance. Therefore, the hair cell τ_m can change with V_m. The range of τ_m depends on cell size, the number of MET channels, and their open probability which depends on intracellular Ca^{2+} buffering and determines the symmetry of $I_{Tr}(t)$, the resting potential and the numbers and types of K^+ channels open at rest. Differences in these parameters result in differences in the receptor potentials of IHCs and OHCs, as discussed below, and may strongly influence tonotopic differences and interspecies differences. The hair cell τ_m may largely determine the frequency ranges of tuning, mechanical feedback and phase locking in the auditory nerve. These electrical properties give rise to two forms of the receptor potential of varying proportion in different hair cells, one that is synchronized to the fine-structure of the sound waveform (the AC potential) and one that is proportional to the envelope of the sound waveform (the DC potential; Russell & Sellick, 1978).

Outer hair cells. In addition to the single row of ∼3000 IHCs, the human cochlea has ∼12,000 outer hair cells (OHCs) arranged in three rows. The silent current, which can be measured extracellularly as a gross cochlear potential, originates primarily from the OHCs (Zidanic & Brownell, 1990) because they are more numerous and have larger MET currents (Fettiplace & Kim, 2014). Although OHCs have no direct influence on the silent current of IHCs or the spontaneous activity of ANFs driven by IHCs

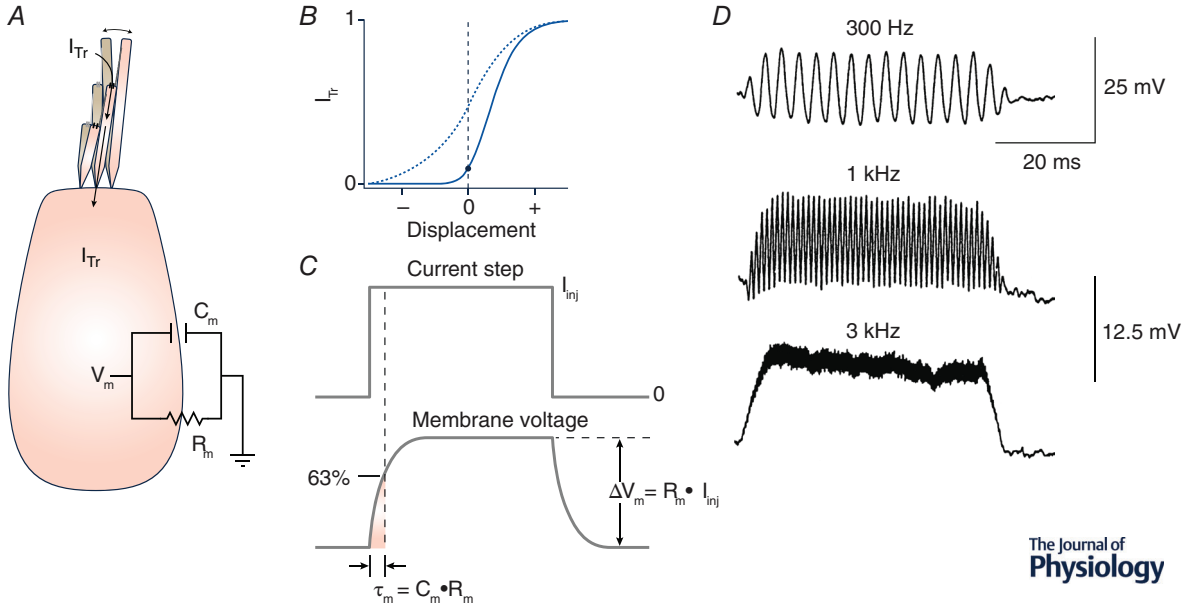


Figure 2. Transduction, membrane time constant and frequency dependence of the hair cell receptor potential
A, hair cell cartoon showing the path of mechano-electric-transduction current (I_{Tr}) through the hair bundle to charge the basolateral transmembrane potential (V_m). Circuit diagram of the basolateral membrane with membrane capacitance (C_m) and resistance (R_m) in parallel. *B*, normalized transduction current amplitude (pA) as a function of normalized hair bundle displacement. In the resting position ($d = 0$), I_{Tr} is non-zero; the 'silent current' depolarizes the hair cell at rest to generate the spontaneous ANF spike rate. Displacement in the excitatory direction increases the depolarizing current; displacement in the inhibitory direction decreases the depolarizing current. The dotted blue line is the symmetric current–displacement function observed for OHCs when the hair bundle is bathed in endolymph (low Ca^{2+}, high K^+). The continuous blue line is the asymmetric current–displacement function observed when the OHC bundle is bathed in perilymph (high Ca^{2+}, low K^+). In endolymph, IHCs from the apical cochlea may have current–displacement functions similar to the dotted line, while those of IHCs from the basal cochlea are similar to the continuous line (Johnson *et al*. 2011; Johnson, 2015). *C*, a square pulse of depolarizing current is injected to the hair cell and the membrane potential is recorded. The hair cell membrane time constant (τ_m) limits the rate of change in V_m as current flows first to charge the membrane capacitance (C_m) and then to cross the membrane resistance (R_m). $\tau_m = R_m \times C_m$, the time to charge the membrane to 63% of the total voltage change. The total change in transmembrane potential, or depolarization (ΔV_m), is the product of the injected transmembrane current and resistance, $\Delta V_m = I_{inj} \times R_m$. *D*, *in vivo* intracellular recordings of IHC receptor potentials in response to tones of different frequencies, as indicated by Russell & Sellick (1983). The responses contain alternating components (AC) that synchronize to the tone stimulus and steady or 'direct' components (DC). At 300 Hz the response is purely AC. As frequency increases, the AC component is filtered out due to the membrane time constant. The DC component is depolarizing due to asymmetry of the IHC current–displacement function. The intensity of the sound was 80 dB SPL for all frequencies. The resting potential of the IHC was −37 mV.

(Dallos & Harris, 1978), OHCs are responsible for hearing sensitivity and sharpness of tuning as summarized below.

The frequency selectivity of cochlear output is measured by the frequency-tuning curve of single ANFs, which are V-shaped functions of response threshold (in decibels SPL) *vs*. sound frequency (Kiang *et al*. 1965). At the tip of the curve, they have a minimum response threshold at the characteristic frequency (CF) of that ANF. At a given cochlear position or CF, the basilar membrane vibration has a frequency-tuning curve with the same shape as that of the IHC receptor potential and the ANF discharge (Sellick *et al*. 1982; Narayan *et al*. 1998). In the absence of the OHCs, frequency tuning is broader and responses at CF are 100 to 1000 times less sensitive (Dallos & Harris, 1978). OHCs are thought to generate this sensitivity and sharpness of tuning through amplification of cochlear vibrations on a cycle-by-cycle basis throughout the frequency range of hearing. In animals that hear at very high frequencies, like mice and bats with ANFs tuned to 70 kHz or more (Russell *et al*. 2003; Taberner & Liberman, 2005), how can OHCs transduce and amplify at such high frequencies? If $f_c = 1/2\pi\tau_m = 70$ kHz, then $\tau_m = R_m \times C_m = \sim 2\ \mu$s. It is not immediately obvious how this could be possible. A hair cell with a reasonable R_m of 50 MΩ would require a seemingly impossibly small C_m of 0.04 pF. A hair cell with a small but reasonable C_m of 4 pF would require a seemingly impossibly small R_m of 0.5 MΩ. Moreover, the actuator of mechanical amplification would need to respond with a limiting time constant of 2 μs to amplify 70 kHz vibrations.

Outer hair cell electromotility is a phenomenon observed in isolated OHCs by which the length of the cell changes with $V_m(t)$ (Brownell *et al*. 1985; Ashmore, 1987). With the mechanical load imposed by neighbouring cells in the tissue, this electromotility manifests as changes in stiffness of the organ of Corti through forces exerted by the piezoelectric motor protein prestin in the OHC basolateral membrane (Zheng *et al*. 2000; Liberman *et al*. 2002; Dallos, 2008). These changes in OHC stiffness provide cycle-by-cycle amplification of cochlear vibrations preferentially for low-level sounds, increasing sensitivity and sharpening cochlear frequency-tuning (Fettiplace, 2017). Indeed, isolated OHCs generate force at frequencies up to 80 kHz (Frank *et al*. 1999), which may produce precise, rapid, high-frequency responses that overcome viscous dampening. In other experiments, the speed limit of OHC motility *in vivo* has been measured at a corner frequency around 3 kHz (Vavakou *et al*. 2019), and thus the presence of a prestin-based mechanism for mechanical amplification at high CF remains controversial. Nonetheless, faster membrane time constants would enable $V_m(t)$ to change faster, which may increase the cycle-by-cycle frequency limit of mechanical amplification by OHCs and/or increase the limit of periodicity encoding by IHCs. For prestin-based amplification to work, the OHC R_m and C_m must be small. Several lines of observation suggest that hair cell properties are indeed tuned to support high-frequency drive to the voltage-dependent stiffness changes exerted by prestin. The expression of $I_{K,n}$ lowers R_m, and coincides developmentally with the onset of electromotility of the cell body of isolated OHCs (Marcotti & Kros, 1999). Relative to IHCs, OHCs have high levels of cytoplasmic Ca^{2+} buffers, which increases MET open probability (Hackney *et al*. 2005). Because the OHC MET channels are half-open at rest, the current–displacement function is symmetric (Fig. 2*B*). Thus, a sinusoidal displacement of the hair bundle will evoke a corresponding sinusoidal voltage excursion around the resting potential (AC) and no summating potential (DC). From the low-frequency cochlear apex to the high-frequency base, there is an increase in MET current and K^+ current (Mammano & Ashmore, 1996; Ricci *et al*. 2003) and a decrease in C_m, all of which shorten the membrane time constant. The OHC τ_m was predicted at body temperature, based on extrapolations from *ex vivo* measurements at room temperature. With increasing CF from apex to base, the OHC τ_m predicted at body temperature range from ∼0.6 ms at the 0.35 kHz location to ∼0.025 ms at the 10 kHz location (Johnson *et al*. 2011; Fettiplace, 2017). τ_m of 25 μs corresponds to a f_c of 6.4 kHz, suggesting that such minimal τ_m filtering *in vivo* may not impede activation of prestin.

In addition to prestin-based somatic electromotility, a second force generator may reside in the hair bundle, the MET channel itself (Kennedy *et al*. 2005; Hudspeth, 2014). The MET channel kinetics have also been proposed to function in cochlear frequency tuning by acting as a band-pass filter, in which the kinetics of MET channel activation is the low-pass filter and MET channel adaptation is the high-pass filter. Indeed, the adaptation time constants of MET channels were faster in rat OHCs tuned to higher frequencies (Ricci *et al*. 2005), consistent with the high-pass mechanism. The hypothetical low-pass mechanism is simply the channel activation kinetics, since the channel requires some finite amount of time to detect the stimulus and change its conformation from closed to open. In this scenario, higher-frequency stimulation would be filtered out because the channel may not open before the stimulus heads towards the trough of the sine wave. However, the activation kinetics of MET currents of mammalian OHCs has so far proved impossible to measure. If the kinetics of MET activation are not to limit the frequency range of mammalian hearing (up to ∼100 kHz in bats, Russell *et al*. 2003), then the activation kinetics of the MET channel must be faster than 10 μs in OHCs of the cochlear base. Such fast kinetics, if present, have so far eluded detection for two reasons. First, the time course for hair bundle displacement cannot be reduced

below ∼50 μs with the use of conventional piezoelectric stimulators (Ricci *et al*. 2005), although new probes have recently been developed that can produce step deflections in less than 10 μs (Doll *et al*. 2012). Secondly, assuming the bundle stimulation was fast enough to stimulate such fast MET channel kinetics, the whole-cell patch-clamp recording bandwidth needs to be fast enough to faithfully record the current. For now, the presence of a MET channel bandpass filter to mediate bandpass frequency tuning along the tonotopic axis remains a hypothesis to be tested.

Inner hair cells. In the mature auditory nerve, both spontaneous and sound-evoked spikes depend on Ca^{2+}-dependent synaptic exocytosis from IHCs, which may be driven by the IHC silent current that is dependent on the endocochlear potential generated by the stria vascularis (Robertson & Paki, 2002; Sewell, 1984). Although the pattern of spontaneous spikes may appear to be random, recent work has revealed negative serial correlations between successive inter-spike intervals in ANFs, which may reflect dynamics of synaptic vesicle pool depletion and replenishment in IHCs (Peterson *et al*. 2014; Heil & Peterson, 2017). Here we consider the activity evoked by sound. Will the ANF fire a spike in response to a given stimulus to the IHC and, if so, when will it spike? How do limitations in the size and speed of changes in the IHC receptor potential influence response threshold and the upper limits of phase synchrony?

Hair cells of non-mammalian vertebrates are frequency tuned by electrical resonance mediated by interaction of voltage-gated Ca^{2+} influx through $Ca_V1.3$ channels and Ca^{2+}-activated K^+ efflux through $K_{Ca}1.1$ channels (Fettiplace & Fuchs, 1999). Prestin-based electromotility and the presence of OHCs is unique to the mammalian cochlea, where frequency tuning is mediated by the OHC-based mechanism of amplification discussed above. In the cochlea, IHCs are not intrinsically frequency tuned; their frequency selectivity depends on their position along the tonotopic axis, where OHCs provide frequency-specific mechanical amplification to IHC mechanotransduction through movement of endolymph. Mammalian hair cells employ $Ca_V1.3$ and $K_{Ca}1.1$ but not for electrical resonance (Brandt *et al*. 2003; Housley and Ashmore, 1992; Dulon *et al*. 1995).

Relative to OHCs, IHCs have less intracellular Ca^{2+} buffer and thus lower MET open probability, a smaller silent current, asymmetric MET current, more hyperpolarized resting potentials (∼−55 mV for IHCs *vs*. ∼−35 mV for OHCs), smaller K^+ conductance at rest, greater C_m and thus longer membrane time constants (Fig. 2*B*; Johnson *et al*. 2011). Thus, the receptor potential in IHCs is asymmetric, with a larger depolarization on one half-cycle than hyperpolarization on the other half-cycle. Between the IHC receptor current $I_{Tr}(t)$ and the receptor potential $V_m(t)$, high frequencies are impeded by the time required to charge C_m (Fig. 2*C*). In other words, gradations of $V_m(t)$ represent changes in sound pressure $W(t)$, but these gradations are limited by the low-pass filter property of the IHC membrane time constant $\tau_m = R_m \times C_m$, which was estimated to be ∼0.2–1 ms from *in vivo* recordings using sharp electrodes in guinea pig (Russell & Sellick, 1978). Similar estimates were obtained in *ex vivo* recordings at body temperature from gerbil IHCs with patch pipettes (Johnson, 2015). Experiments in isolated IHCs from guinea pig at body temperature found K^+ currents with activation time constant of 0.15–0.35 ms that may underlie τ_m and therefore determine f_c (Kros & Crawford, 1990). Changes in $W(t)$ slower than the corner frequency f_c modulate $V_m(t)$ in synchrony with the sound waveform on a cycle-by-cycle basis. Due to the asymmetry of IHC MET, the increase in current for excitatory deflections of the bundle is larger than the decrease in current for inhibitory deflections (Fig. 2*B*). Changes in $W(t)$ faster than the corner frequency, $f_c = 1/2\pi\tau$ (160–800 Hz), are attenuated in $V_m(t)$ as the AC component of $V_m(t)$ is diminished and replaced by the DC component (Fig. 2*D*; Russell & Sellick 1983). At high sound frequency, changes in $V_m(t)$ reflect the temporal envelope of $W(t)$ rather than its fine structure, i.e. by AM rather than by oscillation at the carrier frequency. The asymmetry of transduction in IHCs enhances the DC form of the receptor potential (the summating potential), with direct effects on the temporal pattern of synaptic transmission and ANF discharge, which is phase-locked to the cycles of the acoustic stimulus only at low frequencies (Crawford & Fettiplace, 1980). In addition to these differences between OHC and IHCs, the electrophysiological properties of IHCs may be intrinsically specialized tonotopically to encode the phasic or sustained components of receptor potentials. IHCs from the low-frequency cochlear apex of gerbils had significantly more depolarized resting membrane potentials, faster kinetics and shorter membrane time constants than IHCs from the high-frequency cochlear base. The faster kinetics of low-CF IHCs may allow them to follow the phasic components of sound while high-CF IHCs encode the intensity of high-frequency sounds (Johnson, 2015).

Palmer & Russell (1986) argued that the IHC τ_m limits the frequency of evoked periodic activity in ANFs by showing that τ_m can be derived from estimates of the corner frequency for modulation of ANF spike rate by periodic stimuli. Assuming $\tau_m = 0.2$ ms and an IHC $C_m = 10$ pF results in $R_m = 20$ MΩ. This value is in good agreement with estimates of 20–40 MΩ for *in vivo* sharp electrode measurements of R_m in IHCs (Dallos, 1985). Such a small R_m will speed τ_m, but it creates an apparent sensitivity problem for the IHC. For example,

if a soft sound generates a small MET of I_{Tr} = 10 pA, a R_m = 20 MΩ produces a depolarization of only 0.2 mV ($\Delta V = R_m \times I_{Tr}$). How can such a small voltage change be transduced into a change in spike rate or timing in the ANF? The answer may be related to the operating point of the IHC–ANF synapse at rest. Due to the silent MET current in the absence of sound (Fig. 2*B*), the IHC is depolarized enough to open some $Ca_V1.3$ channels to evoke glutamate release, generating the ANF spontaneous rate. Rather than turning the system on, sound modulates the on-going activity; the IHC at rest need not cross a discontinuity in the input/output function in order to signal the presence of a stimulus. Indeed, one manifestation of this phenomenon is that ANFs begin to synchronize their spikes to the periodicity of low-frequency tones at 10–20 dB below the SPL required to increase the mean spike rate above the spontaneous rate (Johnson, 1980).

The hair cell receptor potential $V_m(t)$ is converted into a spike probability density $P_s(t)$ in each of its primary afferent neurons through mechanisms of synaptic transmission and spike generation. Between $V_m(t)$ and $P_s(t)$ the ribbon synapse and spike generator perform an analog to digital transformation. In general, the instantaneous probability of release from the IHC (P_r) and spike rate in the ANF $P_s(t)$ are proportional to $V_m(t)$. However, the stochastic mechanisms between $V_m(t)$ and $P_s(t)$ introduce jitter and an additional delay of approximately 1 ms (Fig. 3). These steps include opening of voltage-gated Ca^{2+} channels, Ca^{2+} influx and diffusion, binding of Ca^{2+} to the voltage sensor for exocytosis, fusion of synaptic vesicle membrane with the IHC plasmalemma, diffusion and binding of glutamate in the cleft to receptors on ANFs, opening of ionotropic AMPA receptors and depolarization of the heminode to spike threshold.

Synaptic transmission

Threshold responses. Gradations in $V_m(t)$ modulate the voltage-gated calcium current (I_{Ca}) controlling probability of glutamate release $P_r(t)$ from the IHC ribbon synapse. Although still an active area of research with outstanding questions, as discussed above, responses to sound near behavioural threshold levels seem to depend on positive mechanical feedback by active processes in the OHCs to overcome viscous dampening of the acoustic energy (Hudspeth, 2014). Remarkably, at the threshold of hearing, mechanical displacements of the organ of Corti and the IHC stereocilia are smaller than 1 nm (one-billionth of a metre; Sellick *et al.* 1982; Chen *et al.* 2011). Excitability near threshold also depends on the large mammalian endocochlear potential, which is about +80 mV (Davis, 1965; Wang *et al.* 2019). The difference between the endocochlear potential and the IHC resting potential is the electrochemical driving force for I_{Tr}. Assuming an *in vivo* resting potential around −40 mV to −50 mV (Dallos, 1985; Palmer & Russell, 1986), the driving force of ∼120 mV makes relatively large MET current for relatively small bundle displacements. The endocochlear potential thus boosts the amplitude of the MET current, I_{Tr}. Non-mammalian species have smaller electrochemical driving forces for I_{Tr}, and do not hear low-level sounds as well as mammals (Schmidt & Fernández, 1962; Köppl & Manley, 2019). The specialized middle ear bones of mammals also increase sensitivity at low sound levels. Still, as estimated above, near-threshold sounds may depolarize $V_m(t)$ by less than 1 mV, much smaller than the action potential to trigger release at conventional synapses between neurons.

The IHC has unusual mechanisms to support evoked synaptic transmission in response to such small changes in $V_m(t)$ as occurs for near-threshold stimulation (Ashmore, 2010; Rutherford & Pangršič, 2012). For Ca^{2+}-dependent transmission, there must be an increase in the probability of release P_r evoked by the stimulus. The predominant voltage-gated Ca^{2+} current in hair cells is an L-type channel $Ca_V1.3$ (Platzer *et al.* 2000; Brandt *et al.* 2003). L-type Ca^{2+} channels are known for being activated only by 'high' voltage, but in hair cells $Ca_V1.3$ channels begin to open at 'low' voltages around −70 mV to trigger exocytosis (Engel *et al.* 2002; Keen and Hudspeth, 2006; Zampini *et al.* 2010; Graydon *et al.* 2011). The channels open at 'low' voltage (i.e. more negative, more hyperpolarized potential), when the electrochemical driving force for the Ca^{2+} current driving exocytosis is relatively large. Opening at low voltage, the Ca^{2+} influx also happens with minimal delay compared with neuronal synapses that use high voltage-activated Ca^{2+} channels to drive exocytosis with action potentials. Each ribbon synapse in the mouse IHC is estimated to contain 20–330 $Ca_V1.3$ channels packed into a cluster of ∼70 nm × 100–600 nm in size, flanked by docked vesicles, to comprise the active zone (Frank *et al.* 2010; Rutherford, 2015; Neef *et al.* 2018). The fraction of channels available to open at any given time is increased by the effects of calcium binding proteins (CBPs) in IHCs that are thought to limit Ca^{2+} and calmodulin-dependent inactivation of I_{Ca} (Johnson & Marcotti, 2008; Grant & Fuchs, 2008; Hardie & Lee, 2016; Picher *et al.* 2017*a*). Limited inactivation allows I_{Ca} to be sustained with the duration of the depolarization (Fig. 3*A*). As well, Rab interacting molecule (RIM) and RIM-binding protein (RBP) enhance I_{Ca}, either by limiting channel inactivation (RIM2α: Gebhart *et al.* 2010; RBP2: Ortner *et al.* 2020) or by promoting trafficking of $Ca_V1.3$ to the AZ (RIM2α, RIM2β and RIM3γ: Jung *et al.* 2015; Picher *et al.* 2017*b*; RBP2: Krinner *et al.* 2017). Also required for hearing is the Cavβ2 subunit, which regulates Cav1.3 abundance and clustering at presynaptic AZs (Neef *et al.* 2009) along with endophilin-A (Kroll *et al.* 2019).

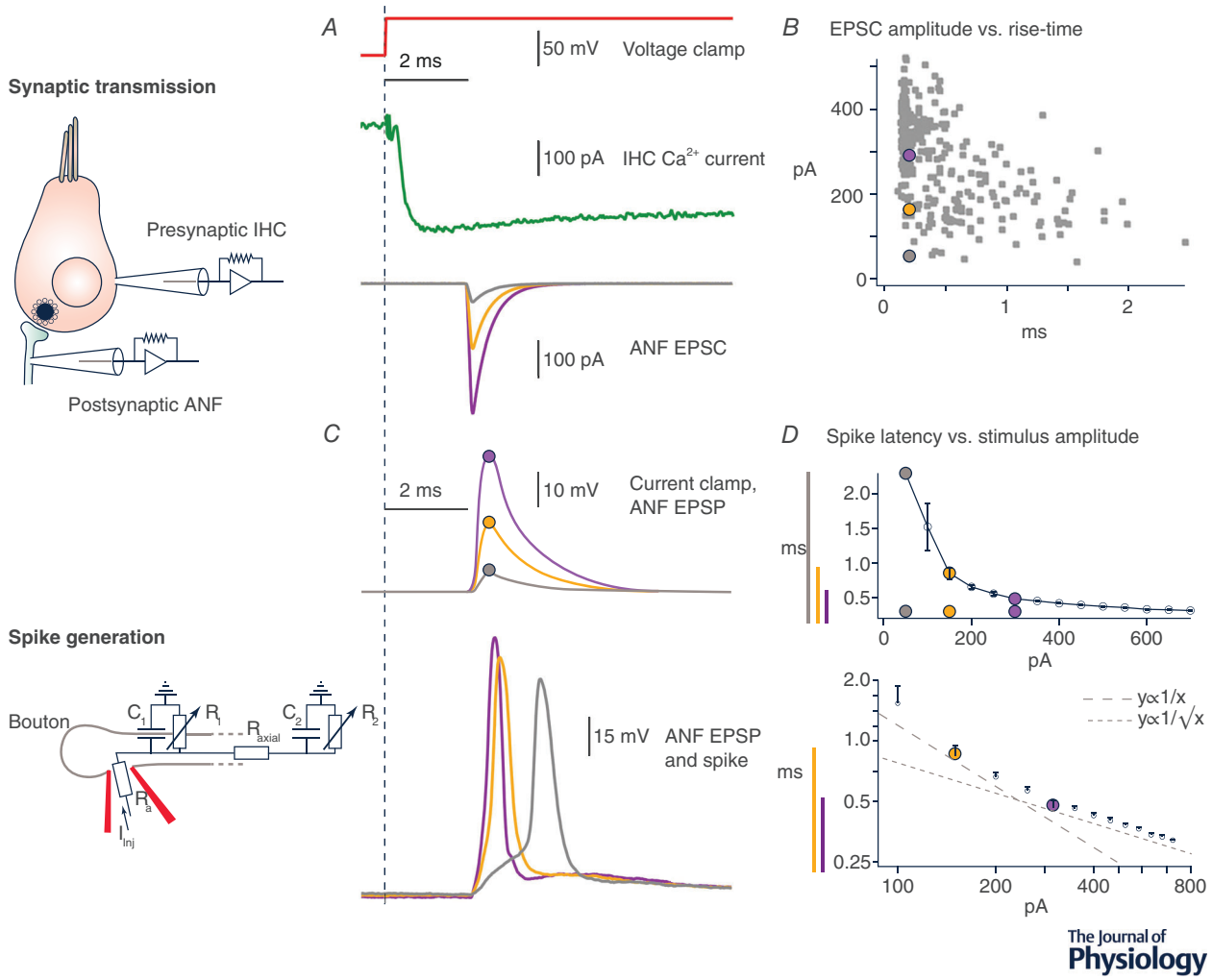


Figure 3. Latency of synaptic transmission and spike generation
A, cartoon depiction of an experiment in the excised organ of Corti, utilizing paired pre- and post-synaptic whole-cell recordings to measure transmission at the ribbon synapse. A step depolarization (red trace) evokes a presynaptic Ca^{2+} current in the IHC (green trace). When the depolarizing step is repeated three times, the resulting presynaptic Ca^{2+} current is very reproducible (shown only once, for clarity). Bottom traces show the first excitatory postsynaptic current (EPSC) in the ANF in response to three identical depolarizing pulses. Traces are representative of data from a young rat, at room temperature, showing a long synaptic delay of approximately 2 ms, on average (Goutman & Glowatzki, 2011). Due to the stochasticity of presynaptic release, the EPSC latency will vary (not shown) and the EPSC size may range from tens of pA to hundreds of pA in response to identical presynaptic Ca^{2+} current. *B*, stochasticity of presynaptic release reflected in the heterogeneity of EPSC peak amplitude (*y*-axis) and 10–90% rise-time (*x*-axis). Data are from one postsynaptic recording of a train of spontaneous EPSCs (i.e. no stimulus applied to the presynaptic hair cell). The grey-, orange- and purple-filled circles mark the parameter space of the three EPSCs, with corresponding colour, shown in *A* and simulated below. *C*, cartoon depiction of an experiment in the same preparation, utilizing a post-synaptic recording to measure spike generation. Synaptic transmission has been blocked with the glutamate receptor antagonist 6-cyano-7-nitro-quinoxaline-2,3-dione. In current-clamp mode, simulated synaptic current is injected (I_{inj}) via the recoding pipette (red, which has an access resistance, R_a) directly into the ANF, which was modelled as two *RC* compartments in series, separated by an axial resistance (R_{axial}). Upper traces, the simulated synaptic current injections from the parameter space indicated in colour in *B* produce the grey, orange, and purple excitatory postsynaptic potentials (EPSPs). Lower traces, larger EPSPs evoke spikes with briefer latencies. The traces in panels *A* and *C* have the same time scale. *D*, same experiment as in *C*. Spike-onset latency (ms) as a function of stimulus (I_{inj}) peak amplitude (pA). When stimuli were >300 pA, latencies were <0.5 ms and temporal jitter (SD of latency) was <20 μs. The grey-, orange-, and purple-filled circles mark the amplitudes of the three stimuli; the vertical lines of corresponding colour mark the spike-onset latencies for the three stimuli. Upper graph is on a linear scale; lower graph is on a log-log scale. As simulated synaptic currents (*x*-axis) approach 300 pA, spike-onset latencies (*y*-axis) become much faster ($y = \sim 1/x$), large-dashed line. Above 300 pA, spike-onset latencies change relatively little ($y = \sim 1/\sqrt{x}$), small-dashed line (Rutherford *et al*. 2012).

Stimulus–secretion coupling. To maintain rapid transmission without depletion of vesicle pools, hair cells are thought to enable indefatigable signalling by virtue of many synaptic vesicles cycling through different morphological pools, to finally enter the readily releasable pool for exocytosis, followed by endocytosis (Moser & Beutner, 2000; Beutner & Moser, 2001; Rutherford & Roberts, 2006; Schnee *et al.* 2005; Johnson *et al.* 2005, 2008, 2009; Kroll *et al.* 2019; Tertrais *et al.* 2019). Together with the close positioning of synaptic vesicles with Ca^{2+} channels (nanodomain coupling, Brandt *et al.* 2005; Graydon *et al.* 2011; Wong *et al.* 2014; Johnson *et al.* 2017), these properties of the IHC AZ enable modulation of release probability for relatively small changes in $V_m(t)$. The upper frequency limit for modulation of release, and thus for modulation of spike-timing in ANFs, can be accounted for by τ_m of the IHC if the kinetics of I_{Ca} are not limiting. Sinusoidal modulation of I_{Ca} by $V_m(t)$ declines with increasing frequency, but f_c for I_{Ca} (∼900 Hz) is higher than f_c for $V_m(t)$ (∼500 Hz) in mouse IHCs *ex vivo*, indeed, suggesting that the kinetics of presynaptic I_{Ca} activation in the IHC are not limiting modulation of the ANF spike rate $P_s(t)$ *in vivo* (Li *et al.* 2014; Wong *et al.* 2013).

Single-channel recording of native Cav1.3 channels in IHCs revealed very brief latency to opening with depolarization (∼50–200 μs), sufficiently rapid to transmit fast signals at sound onset and to allow for high-frequency tuning and phase locking to sound (Zampini *et al.* 2013, 2014). Moreover, Cav1.3 channel opening is observed at membrane potentials as low as −70 mV (Zampini *et al.* 2010), and thus a fraction of these channels are open at the IHC resting potential to drive glutamate release and spontaneous spiking in ANFs. These recordings also showed that the maximal open probability is relatively low (0.1–0.3), and only ∼0.01 at the IHC resting potential, corresponding to ∼10 ms of channel open time, on average, per second. Taking into account that the mean open time for Cav1.3 is <1 ms, the channel opening rate near the resting potential would not be less than 10 s^{-1}. This may imply that the 100 spike s^{-1} spontaneous rate observed in some mammalian ANFs (Kiang *et al.* 1965) can be supported by very few Cav1.3 channels, assuming that the opening of one or very few channels is sufficient to evoke exocytosis (i.e. nanodomain coupling between Cav1.3 channels and synaptic vesicles). The alternative to nanodomain coupling is microdomain coupling, a scenario in which the Ca^{2+} influx from many channels summates to reach the concentration required to evoke the fusion reaction. There may be tonotopic variation in the coupling between Cav1.3 and synaptic vesicles, but even in those IHC synapses that may not be controlled by a nanodomain mechanism (Johnson *et al.* 2008, 2017), the number of Ca^{2+} channels present at each synapse (20–330; Neef *et al.* 2018) would be enough to sustain high spontaneous spike rates at the IHC resting potential.

Synaptic release kinetics depend on the resting state of the synapse, stimulus history and stimulus strength. Goutman & Glowatzki (2011) employed paired IHC–ANF recordings and applied different holding potentials to the IHC before stimulation. They found that short-term Ca^{2+}-dependent facilitation modulated the size and timing of the synaptic response. From a holding potential of −89 mV, brief 10 ms depolarizing pulses to the IHC elicited excitatory postsynaptic currents (EPSCs) in ANFs only approximately 50% of the time. However, when holding potentials were depolarized into range of the *in vivo* resting potential of approximately −50 mV, the test pulse success rate increased to 100%, the size of the postsynaptic response more than doubled (from ∼100 to ∼300 pA), and synaptic latency was halved (from ∼3 to ∼1.5 ms). At the holding potential of −50 mV, a small Ca^{2+} current enters the cell, producing facilitation of synaptic transmission resembling Ca^{2+}-dependent facilitation found in central synapses. These experiments suggest that the synaptic release machinery of IHCs rests in a continuously facilitated state, making the postsynaptic response faster and larger (Goutman & Glowatzki, 2011). Similar short-term facilitation was also observed at the frog auditory synapse with paired recordings (Cho *et al.* 2011; Chen & von Gersdorff, 2019). In addition to triggering exocytosis, presynaptic Ca^{2+} also influences the size of the releasable vesicle pool through vesicle mobilization and priming (Schnee *et al.* 2005, 2011; Cho *et al.* 2011; Levic *et al.* 2011; Castellano-Muñoz *et al.* 2014). We will consider Ca^{2+}-dependent synaptic facilitation again, below, in the section 'Response phase and dependence on stimulus level'.

One interesting feature of IHC–ANF transmission in mammals is that the amplitude of individual EPSCs may not correlate with the level of IHC depolarization and Ca^{2+} influx at steady state (Glowatzki & Fuchs, 2002; Goutman & Glowatzki, 2007). Individual EPSCs come in a wide range of amplitudes (Fig. 3*B*), and while the rate of EPSCs increases with the level of a steady step depolarization, average amplitudes do not change when $V_m(t)$ is varied within the physiologically relevant range of potentials (∼−50 to −20 mV). This may mean that the postsynaptic drive to trigger an individual spike does not change with the level of the IHC receptor potential; or in other words, with sound level. If it did, louder sounds might produce larger EPSCs that would accelerate spike generation and produce phase shifts (Fig. 3*B* and *D*; Rutherford *et al.* 2012). We note that when frog or turtle hair cells are held at very negative potentials (∼−90 mV), EPSCs do become smaller in amplitude (Li *et al.* 2009; Schnee *et al.* 2013), suggesting a Ca^{2+}-dependence to EPSC amplitude.

Temperature. One important and relatively understudied factor is the dependence of cellular mechanisms on temperature. *Ex vivo* experiments in mammalian tissue may be performed at room temperature to prolong the viability of the preparation. Some groups routinely perform their experiments near body temperature (Johnson *et al.* 2005) and others typically do not (Moser & Beutner, 2000). However, other methodological differences hinder comparisons between research groups, and studies directly assessing temperature effects at the IHC–ANF synapse are relatively uncommon. Nonetheless, *in vivo* experiments show that temperature is an important variable that should be carefully considered going forward. For example, moderate cooling of the guinea pig cochlea by just a few degrees below body temperature (38°C) produced reversible effects on the compound action potential (CAP) of the auditory nerve, specifically an increase in latency and an elevation of threshold (Brown *et al.* 1983). These experiments also revealed a decrease in sensitivity of gross cochlear hair cell receptor potentials measured by the cochlear microphonic (CM) at the round window. Interestingly, changes in the CAP and the CM were specific to responses to high-frequency tones, suggesting that mechanisms of high CF hair cells may be specifically affected by cooling. Subsequent experiments used gerbils to measure cochlear potentials from different CF regions (Ohlemiller & Siegel, 1992). Changes in the CAP and the CM were greater in the base than in the apex, while reductions in the endocochlear potential were uniform throughout the cochlea, suggesting the greater vulnerability to temperature of CAP thresholds in the base were not due to region-specific effects on the stria vascularis. Single unit recordings from ANFs showed that cooling the gerbil cochlea by ∼10°C elevated thresholds at CF by 0–15 dB SPL for fibres with CFs below 8 kHz, and by 10–30 dB SPL for higher CF ANFs (Ohlemiller & Siegel, 1994). Effects of cooling on reducing the ANF spontaneous spike rate were also CF dependent, being <50% for fibres with CF <4 kHz, and >50% for fibres with CF >4 kHz. The upper frequency limit and the degree of synchronization of phase locking were also reduced in response to tone bursts of 0.5–2 kHz (Ohlemiller & Siegel, 1998). What cellular mechanisms might account for these temperature effects? Temperature effects are likely to be complex, affecting multiple processes, some of which may be specific to high- *vs.* low-frequency hearing, while others may be more general.

In the relatively few published studies comparing results at room temperature and body temperature in the explanted organ of Corti, significant effects have been reported. When temperature was raised from 22 to 37°C, the voltage-gated Ca^{2+} currents of IHCs become slightly larger and much faster (Grant & Fuchs, 2008; Marcotti *et al.* 2003*b*; Johnson & Marcotti, 2008). The relatively small increase in IHC I_{Ca} amplitude is accompanied by a much larger increase in synaptic release as the Ca^{2+}-efficiency of exocytosis increases with temperature (Nouvian, 2007; Pangršič et al. 2015). Along with the increase in Ca^{2+} current activation kinetics, the synaptic delay is reduced at body temperature in mammalian vestibular hair cells (Songer & Eatock, 2013). In experiments on rat IHC–ANF synapses at body temperature, where synaptic transmission was studied in isolation from effects on MET, spontaneous EPSC rates approximately doubled (Wu *et al.* 2016). In the bullfrog amphibian papilla, increasing temperature from 24 to 32°C resulted in shorter synaptic delays, increased spike rates due to increased EPSC rates, increased I_{Ca} and Ca^{2+}-efficiency of exocytosis, and increased size of the readily releasable vesicle pool of synaptic vesicles (Chen & von Gersdorff, 2019). Although EPSC amplitude was unchanged, the EPSC rise and decay were faster, and because Ca^{2+} influx and K^+ efflux increased simultaneously, the hair cell resting membrane potential remained stable during temperature increases (Chen & von Gersdorff, 2019). MET currents measured in 1.3 mM Ca^{2+} (perilymph) from OHCs in the gerbil apex increased from ∼0.7 nA at room temperature to ∼1 nA near body temperature. Changing to 0.02 mM Ca^{2+} (endolymph) produced a similar increase at room temperature, while the fraction of current activated at rest increased from 0.08 to 0.43 (Johnson *et al.* 2011). Similar measurements on the voltage-dependent K^+ current gave maximum currents of ∼7 nA at room temperature and ∼14 nA at 36°C. Therefore, recording conditions including temperature and ionic concentrations can have large effects on cellular electrophysiology. More work is needed to determine which processes produce the temperature dependent effects observed *in vivo*.

Auditory nerve

Divergence from IHC receptor potential to ANF spike trains. From each mammalian IHC several ANFs diverge, carrying information to the brain in parallel. As mentioned above, the IHC receptor potential $V_m(t)$ is a continuous signal because mature IHCs do not fire action potentials due to developmental upregulation of $I_{K,f}$ (Kros *et al.* 1998; Beutner and Moser, 2001; Marcotti *et al.* 2003*a*). As such, the information in $V_m(t)$ is not limited by a maximum spike rate because there is no action potential refractory period preventing $V_m(t)$ from continuously signalling changes in sound over time. To convey information about the sound pressure waveform to the brain, $V_m(t)$ diverges to several ANFs, each of which is excited by one ribbon synapse on one IHC (Liberman, 1980). If all these ribbon synapses and the ANFs they excite responded similarly, they would carry redundant information. Rather, ANFs seem to be excited

heterogeneously, with different aspects of the auditory signal diverging along different ANFs. This parallel representation may increase the temporal resolution of information about $V_m(t)$ that is conveyed to the brain. While one ANF is refractory, or still depolarizing toward spike threshold, another ANF may be spiking in response to a stimulus feature. Divergence from one $V_m(t)$ to several spike trains $P_s(t)$ also increases the resolution of sound level representation in the nerve because individual ANFs have different thresholds and dynamic ranges (Kiang *et al*. 1965; Zagaeski *et al*. 1994). Some ANFs respond to very low-level sounds and saturate their spike rates at moderate SPL, while others don't begin increasing their spike rates until sounds are moderately loud. While the whole nerve response changes over an ∼80 dB range, representing four orders of magnitude difference in SPL, most mammalian ANFs have dynamic ranges of <20 dB (Kiang *et al*. 1965; Taberner & Liberman, 2005). Thus, fractionation of the full range by individual ANFs with limited dynamic ranges may enable increased sensitivity to small changes in SPL if high- and low-threshold ANFs project to different cells in the cochlear nucleus (Liberman, 1991, 1993; Ryugo, 2008).

How is range fractionation achieved? Hair cells may decompose $V_m(t)$ into different parallel outputs through ribbon-type AZs having different voltage-dependence of Ca^{2+} influx through $Ca_V1.3$ channels (Frank *et al*. 2009; Ohn *et al*. 2016). Postsynaptic differences in ANF excitability related to glutamate receptor expression or voltage-gated ion channel expression may combine with presynaptic heterogeneity of Ca^{2+} influx to produce mutual response diversity among ANFs excited by one IHC $V_m(t)$. A major focus of on-going research regards how these presynaptic and postsynaptic heterogeneities combine to determine the diverse sound-response properties of ANFs. The individual sources of heterogeneity may be relatively small but combine to produce significant differences. One unanswered question is: do these heterogeneities combine randomly or in a pattern to increase diversity of encoding?

In auditory organs of non-mammalian vertebrates, each afferent fibre collects input from many synaptic ribbons converging from one or more hair cells, with similarity to the anatomy of mammalian vestibular organs (Eatock & Songer, 2011). In the amphibian papilla, a single fibre may receive as many as 100 or more ribbon synapses from several similarly tuned hair cells, with 30–40 of those synapses coming from a single hair cell contacted by a calyx-type fibre terminal (Simmons *et al*. 1992; Graydon *et al*. 2014). As in the mammalian IHC, the range of sound levels encoded by an amphibian papilla hair cell may be extended by having AZs with different release probabilities, P_r. Differences in P_r among AZs in a hair cell may arise from differences in the number of Ca^{2+} channels, their gating behaviour, or physical coupling to vesicles. Collectively, the AZs of one hair cell could have sensitivity to small and large changes in $V_m(t)$. For non-mammalian vertebrates, a single nerve fibre with multiple AZs of different P_r could enable a larger dynamic range relative to mammalian ANFs. Indeed, afferent fibres in the frog have larger dynamic ranges than those of mouse (Christensen-Dalsgaard & Narins, 1993; Christensen-Dalsgaard *et al*. 1998; Taberner & Liberman, 2005). In either anatomical arrangement, high P_r AZs will release vesicles for small membrane potential changes, whereas lower P_r AZs may be excited later in a stimulus train or only for larger changes in $V_m(t)$, possibly via facilitation at higher global calcium levels. If broad dynamic ranges may be obtained within a single fibre contacting a single hair cell, as in the amphibian papilla, what would be the evolutionary advantage of the divergence found in the mammalian cochlea? Perhaps it is related to the potential advantage afforded by independent modulation of ANFs by cochlear efferent neurons. However, mammals may have paid a price for this multiplexing of information in the cochlea. The thin calibre of mammalian ANFs renders them vulnerable to ionic imbalance during periods of very high activity, such as during noise trauma (Puel *et al*. 1998). This osmotic stress can produce axonal fragmentation and permanent neurodegeneration (Kujawa & Liberman, 2009).

Spike generation. During postnatal development of the rodent cochlea, spike-generating heminodes in the osseous spiral lamina (expressing $Na_V1.1$, $Na_V1.6$, $K_V1.1$, $K_V2.2$, $K_V3.1b$, $K_V7.2$ and $K_V7.3$) align at the habenula perforata where the distal myelin begins (Hossain, 2005; Smith *et al*. 2015; Kim & Rutherford, 2016; Wan & Corfas, 2017). The ion channels of spike generation in ANFs are tightly coupled, both electrically and morphologically, to the afferent ribbon synapses only ∼20 μm away. Because of this tight coupling and the lack of endogenous spiking, the timing of spike generation appears to be largely determined by the timing of exocytosis from the IHC, both spontaneously and in response to sound. One property of ANFs ensuring that spike timing reflects synaptic activity rather than internal neuronal dynamics is the phasic response to depolarization. In organ of Corti explants from mature rats, responses to sustained depolarization are rapidly adapting or phasic; they fire a single spike at stimulus onset (Rutherford *et al*. 2012). This firing behaviour has been termed class III excitability, phasic or single spiking (Hodgkin, 1948; Izhikevich, 2007; Prescott *et al*. 2008). Phasic excitability also occurs in principal cells in the auditory brainstem that are involved in sound localization circuits (Scott *et al*. 2005; Portfors & von Gersdorff, 2013; Huguet *et al*. 2017). Thus, high spike rates in ANFs require rapid repolarization of the

ANF in between events of exocytosis, which may be aided by low voltage-activated K_V1 channels and dendritic HCN channels (Yi *et al.* 2010), also found in neurons of the ventral cochlear nucleus (Oertel *et al.* 2008) and in vestibular primary afferent neurons (Lysakowski *et al.* 2011; Ventura & Kalluri, 2019). Local and tonotopic variation in K_V1 and HCN channels may be a primary mechanism of ANF response heterogeneity (Liu *et al.* 2014a,b; Boulet & Bruce, 2017).

ANFs have a spike threshold in terms not only of amplitude but also the rate of current injection or depolarization, dV/dt. This type of excitability is characteristic of class III neurons, which do not respond to slow stimuli and thereby act as high-pass filters (McGinley & Oertel, 2006; Gai *et al.* 2009). These properties might prevent multiple spikes during longer EPSPs, and thereby enhance the locking of spike times to the onsets of neurotransmitter release events. The phasic property of ANFs may be due in part to Na^+ channel inactivation (Santos-Sacchi, 1993). Curiously, the principal Na^+ channel isoform located at axon initial segments and nodes of Ranvier in ANFs, $Na_V1.6$, is known in Purkinje neurons to be relatively resistant to inactivation and seems to promote repetitive firing (Raman *et al.* 1997). Resistance to Na^+ channel inactivation can enable repetitive firing to a single depolarization via persistent and resurgent I_{Na} that depends on the regulatory subunit $Na_V\beta4$, as seen in cultured spiral ganglion neurons and in auditory brainstem neurons in slices (Kim *et al.* 2010; Browne *et al.* 2017). However, in mature ANFs, other channels that promote hyperpolarization seem to prevent repetitive spiking unless evoked by a sequence of individual stimuli. Together with persistent and resurgent I_{Na} to underlie high rates of depolarization, several types of K^+ channel found on ANF axons may maintain phasic excitability at high rates through rapid repolarization, including $K_V1.1$, $K_V1.2$, $K_V2.2$, $K_V3.1b$ and $K_{Na}1$ (Mo *et al.* 2002; Smith *et al.* 2015; Kim and Rutherford, 2016; Reijntjes *et al.* 2019).

In comparison with the process of synaptic Ca^{2+}-dependent exocytosis evoked by $V_m(t)$ in the IHC, the process of spike generation evoked by excitatory postsynaptic potentials (EPSPs) in ANFs appears to be relatively fast (briefer latency), more reliable (fewer failures) and more precise (less variance in timing; Goutman, 2012; Rutherford *et al.* 2012; Figs 3 and 4). When the IHC releases packets of glutamate, at least at relatively low rates ($<\sim100$ s^{-1}), the EPSP is large and fast relative to voltage threshold and dV/dt threshold. As discussed above, synaptic transmission from hair cells is highly variable in amplitude and wave shape (Glowatzki & Fuchs, 2002; Keen and Hudspeth, 2006; Li *et al.* 2009; Grant *et al.* 2010 ; Schnee *et al.* 2013) (Fig. 3*C*). How is spike generation in ANFs affected by EPSC variability? For spontaneous spikes, the great majority of EPSPs successfully triggered spikes *in vivo*, with only $\sim$12% failure rate (Siegel, 1992). Similarly, in the explanted organ of Corti the rates and interval distributions for EPSCs and spikes were nearly identical (Rutherford *et al.* 2012). The nearly one-to-one conversion of EPSPs to spikes under spontaneous conditions indicates that, in the absence of neural refractoriness, only the smallest release events failed to trigger a spike. At higher rates of synaptic activity (>100 s^{-1}), relative refractoriness may lead to failures, particularly if synaptic depression leads to smaller EPSCs. In auditory endorgans of the turtle and frog or in mammalian vestibular endorgans, where multiple ribbon synapses provide convergent input to each afferent neuron, the success rate for individual spontaneous EPSCs to trigger spikes is thought to be lower than for the cochlea at the IHC–ANF synapse (Schnee *et al.* 2013; Songer & Eatock, 2013; Graydon *et al.* 2014; Li *et al.* 2014). This may be due to the comparatively large input resistance and small capacitance of the mammalian ANF that receives a single ribbon input, compared to the larger frog and turtle calyx-type afferents that receive multiple ribbon inputs and have lower input resistance and greater capacitance (Chen & von Gersdorff, 2019).

In mammalian ANFs the typical EPSC amplitude of approximately 300 pA greatly exceeds spike threshold. When currents were injected through the patch-pipette, EPSC-like waveforms with amplitudes of 50–100 pA were already enough to evoke spikes. Nevertheless, at high rates larger EPSCs may be required to overcome refractoriness. Although small EPSCs triggered spikes, increasing their size to the mean EPSC amplitude dramatically reduced latency and jitter. Spike-onset latencies improved from $\sim$1.5 ms to $<$0.4 ms when increasing the EPSC amplitude from 100 to 300 pA at room temperature, while jitter was reduced from >100 μs to <10 μs (Rutherford *et al.* 2012; Fig. 3*D*). Reduction of spike latency and jitter is the benefit of synaptic events being larger than the minimum required to reach threshold, which is an energetically demanding phenomenon that has been observed across different species. In the amphibian papilla, large EPSCs improved the speed and precision of spike generation and increased phase-locking precision (Li *et al.* 2014).

EPSC and spike latency, time course and jitter. Delays between the external stimulus $W(t)$ and the hair cell receptor potential $V_m(t)$ can be smaller than 1 ms for IHCs of the basal cochlea in response to moderately loud sounds (Kimberley *et al.* 1993). For hair cells in the human cochlear apex, as far as 35 mm from the cochlear base where sounds enter the cochlea, the delay between $W(t)$ and $V_m(t)$ lengthens by a few milliseconds as the traveling wave delay becomes longer. The synaptic ribbon is thought to reduce the latency and increase the temporal precision of EPSCs by increasing the number of synaptic vesicles

readily available for release (Wittig & Parsons, 2008; Buran *et al.* 2010). Nonetheless, perhaps surprisingly, synaptic transmission at the ribbon synapse is the slowest cellular process between sound and afferent discharge, and the largest source of variability that limits the temporal precision of spike timing. The time course of the EPSP and spike measured intracellularly is ∼1 ms. Although the temporal jitter of spike generation may be much less than 1 ms (Fig. 3*D*), the time course and temporal jitter of synaptic transmission may be 1 ms or more (Fig. 4*B*). This temporal jitter is relatively small for low-frequency sounds but becomes large for high-frequency sounds, thus restricting the frequency range of periodicity encoding (Fig. 1*E–F*). Even for hair cells with fast membrane time constants, the temporal jitter of synaptic transmission would seem to limit periodicity encoding to below ∼1 kHz.

The ribbon synapse is also the source of level-dependent differences in sound-response latency, for example at sound onset. Hearing is known as the fastest sense, with latencies between sound onset and first spike in the primary auditory cortex of approximately 12 ms for loud sounds at high CF (mouse: Guo *et al.* 2012). In the cochlear base where the travelling wave delay is very brief, the latency between sound onset and first spike in an ANF can be smaller than 1 or 2 ms. For low-level sounds, first-spike latencies in the same ANF can be several tens of milliseconds (Heil & Irvine, 1997). Loud sounds are heard before soft sounds because perceptual threshold is a function of temporal integration of sound

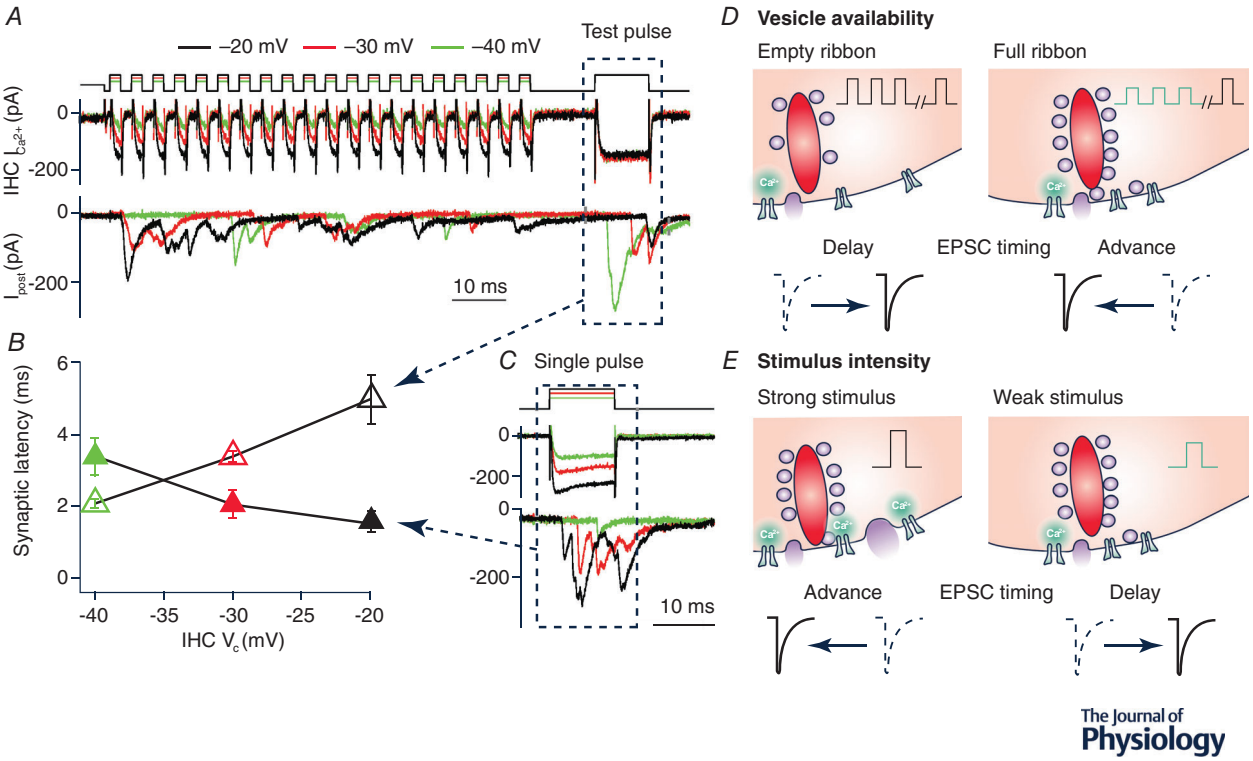


Figure 4. Facilitation and depression may offset to maintain synaptic latency at different stimulus levels
A, paired recording of a ribbon synapse in rat organ of Corti. Top traces show three trains of 2 ms depolarizing steps from the holding potential of −70 mV to −20, −30, or −40 mV. Middle traces show the presynaptic Ca^{2+} current in the IHC. Bottom traces show the EPSC trains in the ANF. At the end of the pulse train, the dashed blue box around the 10 ms pulse to −20 mV is a 'test' to probe for the remaining synaptic vesicle pool. The larger pulse train (black, −20 mV) results in greater depletion, and thus, the response to the test pulse is smaller and longer in latency. The smaller pulse train (green) results in less depletion, and thus the response to the test pulse is larger and faster. *B*, open symbols, synaptic latency in response to the 'test' pulse in *A*. Latency increases as a function of the level of the preceding stimulus train. Filled symbols, synaptic latency decreases as a function of pulse level, in response to the single pulses in *C*. These two effects may offset in the context of an on-going stimulus (Goutman, 2012). *C*, in response to 10 ms depolarizing pulses, larger depolarizations evoked larger Ca^{2+} currents and larger EPSCs with briefer latency (black traces). Smaller depolarizations evoked smaller Ca^{2+} currents and smaller EPSC with longer latency (green traces). *D*, left, in response to a large stimulus train, release from depleted ribbons produces a delay in response to the test pulse. Right, in response to a smaller stimulus train, release from non-depleted ribbons produces an advance in response timing to the test pulse. *E*, left, the larger single pulse will open more presynaptic Ca^{2+} channels, resulting in facilitation which tends toward a phase advance. Right, the smaller single pulse will open fewer channels, tending to delay release.

pressure over time (Heil & Neubauer, 2003). Temporal summation of the pressure envelope appears to happen in the ear because it is present already in the threshold functions of ANFs. This suggests that peripheral latency is the limiting factor, and more central processes are not necessary to explain how perceptual thresholds for sound duration become longer for lower SPL (Heil & Neubauer 2001). As mentioned previously, the IHC membrane time constant is thought to be ∼0.2 ms, considerably shorter than the long integration times observed for first-spike latency in ANFs in response to low-level sounds. Together, these observations strongly suggest that the mechanisms for temporal integration of sound pressure lie between $V_m(t)$ and $P_s(t)$, at the IHC ribbon synapse and/or the ANF spike-generator. Similarly, in the zebrafish lateral line organ, the locus of the dependence of fist-spike latency on stimulus strength was determined to be at the stage of signal transduction between $V_m(t)$ and $P_s(t)$, specifically the hair cell ribbon synapse and/or neuronal spike generator (Troconis *et al*. 2017).

If first-spike latencies closely reflect first-EPSC latencies, as discussed above (Fig. 3*D*) then synaptic latencies at the ribbon synapses in the organ of Corti range from shorter than 1 ms to tens of milliseconds, depending on the stimulus. While synaptic mechanisms between IHCs and ANFs enable response latencies that depend largely on the properties of the stimulus, they also introduce randomness. When presented with repetitions of a moderately loud sound, the first-spike latencies have considerable jitter, ranging from less than 2 ms to greater than 5 ms in some ANFs (Heil & Irvine, 1997; Buran *et al*. 2010). Thus, the latency of synaptic transmission at sound onset seems to depend on the stimulus plus some variable internal state of the synapse. Moreover, a given ANF will not fire at every sound onset in the series of repetitions, due to failures of synaptic transmission (Fig. 1*E*), thought to arise from limited availability of synaptic vesicles. Even in afferent neurons from the saccule, the hearing organ of fish, which receive convergent inputs from active zones on many hair cells, an individual afferent does not respond to every cycle of an on-going stimulus due to vesicle depletion (Furukawa & Matsuura, 1978; Furukawa *et al*. 1982). Similarly, except for perhaps at very low frequencies (<100 Hz), a given ANF will not fire on every cycle of a tone, exhibiting more skipped cycles with increases in frequency and decreases in level (Rose *et al*. 1967). Even if synaptic transmission could match the rate of a high-frequency tone the ANF will skip cycles due to refractoriness of spike generation and the recovery of spike amplitude, both of which limit the ANF spike rate to be below 1 kHz (Peterson & Heil, 2018; Peterson *et al*. 2018). Collectively, the ensemble of ANFs may respond on every cycle, and neurons of the cochlear nucleus receiving converging inputs from the nerve may recover timing information lost by individual ANFs.

If one assumes that spontaneous activity occurs in the absence of an applied stimulus but not during the stimulus, then spontaneous spikes could not be mistaken for first evoked spikes. However, we speculate that spontaneous spikes occurring within about 10 ms in advance of stimulus onset may delay the first evoked spike due to IHC vesicle depletion or ANF refractoriness. If one assumes that spontaneous activity continues during evoked activity, then, for ANFs with low spontaneous rates (<1 spike s^{-1}) the likelihood of a spontaneous spike occurring before the evoked spike is low, and the first spike could be directly observed most of the time. For example, if an ANF has a SR of 1 s^{-1} and if first-spikes occur with ∼10 ms latency, then the spontaneous spike may interfere 1% of the time, on average. For ANFs with high spontaneous rate, the likelihood may be high such that spontaneous spikes occur before evoked ones a significant fraction of the time. In this case one would approximate the mean first spike latency from the peak of the peri-stimulus time histogram. It also depends on the repetition rate of the stimulus, because spontaneous spikes are inhibited for some time after the offset of a strong stimulus (Fig. 1*D*).

Phase locking

Spike timing precision in the ascending auditory pathway. In the auditory system, phase locking refers to the phenomenon whereby spike probability varies as a function of the phase of low-frequency periodic sounds (Fig. 5; Galambos & Davis, 1943; Joris *et al*. 2004; Heil & Peterson, 2015, 2017). For low-frequency sounds, the brain uses interaural timing information present in sound onset, periodicity and stochastic fluctuations to calculate the angle of the horizontal vector to the sound source by comparing the time or the phase of spikes between the two ears (Grothe *et al*. 2010; Ashida & Carr, 2011). The temporal precision for encoding interaural time differences into neuronal spikes in the auditory nerve is limited by the stochasticity of IHC–ANF synaptic transmission, which, as discussed above, can result in millisecond differences in spike timing for repetitions of a given stimulus. Temporal coding, both of fine-structure and of envelopes, is enhanced starting in the cochlear nucleus due to convergence of synaptic inputs, but the upper frequency limit for synchronization decreases as precision increases. For example, the precision of phase synchronization to fine-structure is much enhanced in bushy cells relative to the auditory nerve, but whereas ANFs may show some phase synchrony up to ∼5 kHz, bushy cells do not. This theme is repeated for envelope synchronization, which is much enhanced in octopus cells and many other neuron types in the cochlear nucleus relative to the nerve, but again the highest modulation frequencies for encoding periodicity

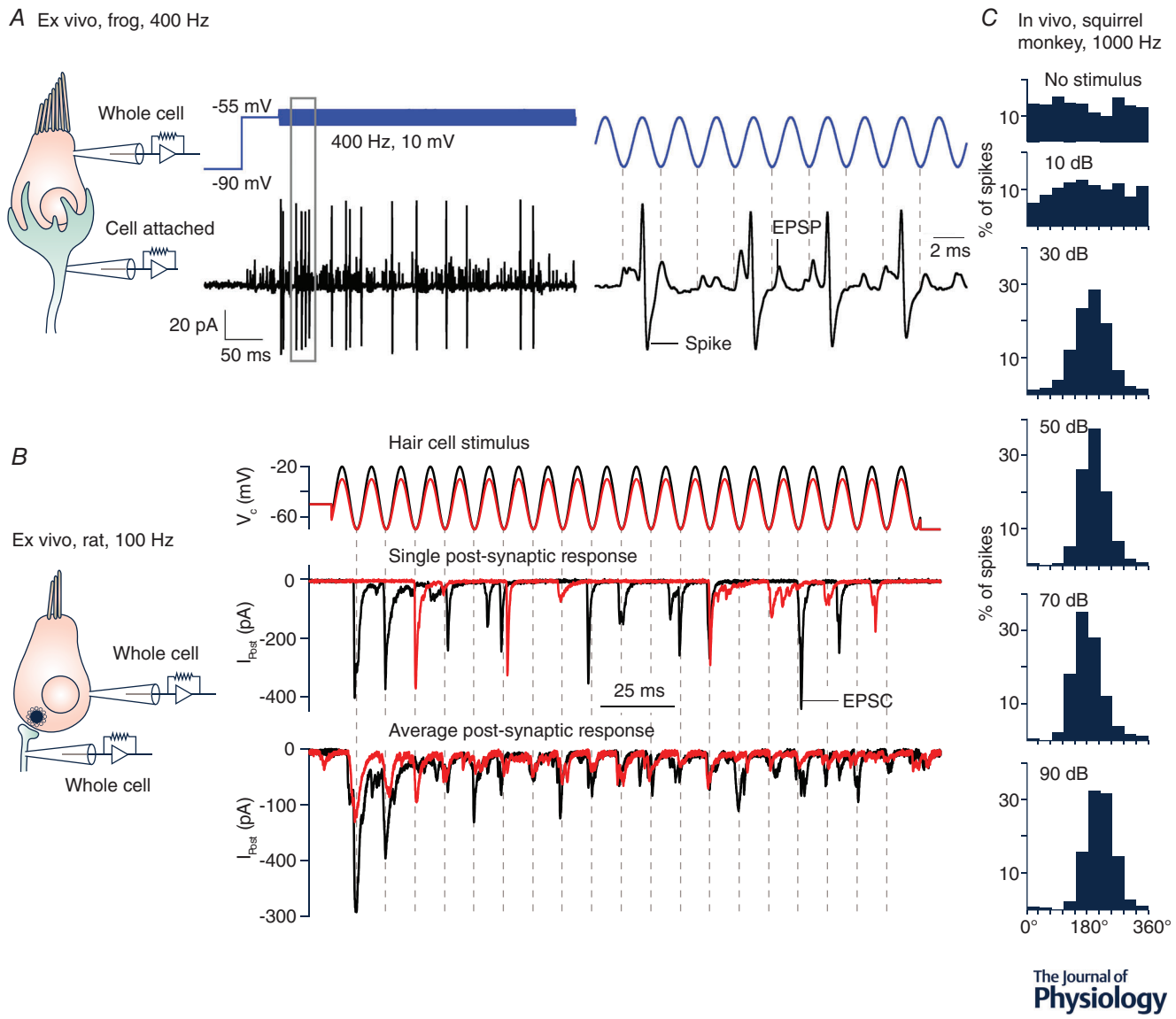


Figure 5. Phase-locking and dependence of response phase on sound level
A, paired recording of phase locking from the bullfrog amphibian papilla. Whole cell voltage-clamp recording from the hair cell, which was held at −90 mV and then at −55 mV, a potential close to the *in vivo* resting potential of a hair cell in silence. A 400 Hz sinewave voltage command was then applied to the hair cell (blue traces). A patch pipette in the cell-attached mode records EPSPs and spikes (black traces). The grey rectangular box area on the left is shown in greater time resolution on the right. Note that many EPSPs do not trigger a spike in the amphibian papilla. Spikes occur during a preferred phase of the sine wave (Li *et al*. 2014). *B*, paired recording of EPSCs in the presence of tetrodotoxin to block spikes at the rat IHC–ANF synapse in response to 100 Hz sinusoidal stimulation at two levels (upper). Responses to the larger stimulus (black) are more frequent, due to increase in the presynaptic Ca^{2+} current (not shown), as shown for two individual responses (middle). Response phase is relatively unchanged with the increase in stimulus level, as can be seen more clearly in the average postsynaptic responses, at the bottom (Goutman, 2012). *C*, period histograms of *in vivo* spike trains from a high spontaneous-rate ANF of the squirrel monkey, showing the phase of occurrence of spike times in the period of the 1 kHz tone stimulus. As SPL increased by 10,000-fold from 10 dB to 90 dB the mean spike rate remained relatively unchanged and the peak in the period histogram remained relatively stationary near 180 degrees, or about half-way through the 1 ms period of the tone (Rose *et al*. 1967).

are smaller than in the nerve (Joris *et al.* 1994*a*,*b*, 2004). The underlying mechanisms of the inverse relationship depend on the context. The cochlear nucleus primarily encodes envelopes temporally. As information ascends the auditory pathway, the code seems to shift away from a temporal code observable in single cells and towards a rate code in larger populations of neurons. In the superior olivary complex, as binaural circuits converge, the inverse relation is apparent as a lower frequency limit for phase-locking while coincidence detection supports encoding of interaural timing differences underlying the behavioural precision of sound source localization.

Remarkably, the lowest behavioural thresholds for sound source discrimination are in the range of 1–2 degrees horizontal. For a low-frequency tone of 1 kHz, these behavioural thresholds rely on detection of interaural timing differences of 10–20 μs (Klumpp & Eady, 1956), i.e. 10–20 millionths of a second. This requires that representation of timing information be not only maintained but significantly improved in the ascending auditory pathway. Given the jitter in the auditory nerve, how is such behavioural precision achieved? Although the answer is not clear, some examples of enhanced precision in the ascending auditory pathway have been well described. ANFs terminate on several classes of cells in the cochlear nucleus. Two of them, the octopus cells and globular bushy cells, are probably the most temporally accurate neurons in the brain, with temporal jitter of several tens to several hundreds of microseconds. Thus, the temporal precision of sound encoding is enhanced from the auditory nerve to the cochlear nucleus, and again from the cochlear nucleus to the trapezoid body containing output axons from the anteroventral cochlear nucleus (Joris *et al.* 1994*a*,*b*). This is achieved in part by anatomical convergence: simultaneous input from many ANFs converge onto one octopus cell to drive it to spike threshold; a few ANFs converge onto each globular bushy cell via the giant endbulbs of Held, each containing hundreds of AZs (Golding *et al.* 1995; Oertel *et al.* 2000; Cao & Oertel, 2010). The ANF inputs can phase lock to higher frequencies but the bushy cell can do so more precisely. The gain in temporal precision may come from convergence of inputs if each ANF spikes at a different time but the bushy cell is made to fire a spike when the first input is activated. Over successive cycles, if the arrival of the first spike across the three ANFs falls in a narrower window than the first spikes in an individual ANF, then the bushy cell will have more precise spike timing. Still, behavioural thresholds are more precise than the temporal jitter of the most precise neurons. In humans, the upper limit of phase-locking to stimulus fine-structure appears to be only ∼1.4 kHz (Joris & Verschooten, 2013), which implies that discrimination of sound source location through encoding of interaural timing differences relies on conversion of a temporal code into a rate code across a population of neurons (McAlpine *et al.* 2001; Salminen *et al.* 2010).

Limitations to phase locking, beyond the hair cell receptor potential. In some ANFs of some species, phase locking to fine-structure or envelope is observed for frequencies up to several kilohertz. For relatively low carrier frequencies, the IHC receptor potential follows the fine-structure of the sound pressure waveform, and for higher carrier frequencies it closely follows the temporal envelope. As introduced above in the section on 'Peripheral spectral and temporal processing', the upper limits on envelope phase-locking are significantly lower than the upper limits on phase locking to fine-structure. Phase locking is thought to be required for localization of low-frequency sounds, for detection of low-frequency sounds near threshold, for pitch perception, and for speech in noise perception (Knudsen & Konishi, 1979; Lorenzi *et al.* 2006; Moore, 2008; Grothe *et al.* 2010; Huet *et al.* 2018). Remarkably, in the barn owl which specializes in using interaural timing differences for prey detection, phase locking in ANFs can occur at up to 9 kHz (Köppl, 1997). Notably, a single afferent fibre receives multiple ribbon synapses in these warm-blooded vertebrates (Köppl, 2011).

In addition to the hair cell membrane time constant, the properties of synaptic transmission and spike generation also limit the upper frequency of phase locking. For frequencies admitted by the hair cell membrane RC time constant τ_m (generally below a few kilohertz), periodic stimuli evoke periodic spikes in ANFs (Fig. 1*E*). A period histogram shows the distribution of spikes over the stimulus cycle. Phase locking can be visualized in period histograms as a non-uniform distribution of spike probability $P_s(t)$ as a function of time within the period (Fig. 5*C*). The degree of phase-locking can be quantified using the measurement of vector strength, between 0 and 1. If most spikes fall within a small range of phase angles, the vector strength is high (Goldberg & Brown, 1969). At moderate sound intensity, frequency tuning is relatively broad, particularly for low-threshold ANFs. Within the frequency limit of phase-locking, if the tone level and frequency are within the response area of the ANF tuning curve, then spikes will synchronize to the tone. For low-frequency tones, the rate of discharge drops to near zero in the valleys of the period histogram, particularly for high-threshold ANFs (Joris & Yin, 1992). For tones between 3 and 5 kHz, $P_s(t)$ does not decrease below the spontaneous rate during any portion of the stimulus cycle and the waveforms of the period histograms are nearly sinusoidal (Johnson, 1980). In guinea pig and cat, phase-locked responses are not observed above 3 or 4 kHz as modulation of $V_m(t)$ vanishes due to the hair cell τ_m. Although weak phase-synchrony to tones can be detected at a few kilohertz, the strength of phase locking

is already significantly diminished at 1 kHz. The temporal jitter of synaptic transmission (Fig. 4*B*), and the influence of stochastic EPSC size on the kinetics of spike generation (Fig. 3) also limit the precision or reproducibility of phase locking and contribute to the decline of synchrony as frequencies increase beyond several hundred hertz.

How does the barn owl extend the upper limit of phase locking to 9 kHz? Addressing the membrane time constant limitation, $\tau_m = R_m C_m$, one consideration is hair cell size. Rather than inner and outer hair cells, birds have tall and short hair cells. Interestingly, the barn owl is different from other birds in that it has a larger proportion of the short hair cells (Fischer *et al*. 1988). Smaller hair cells have faster time constants, extending the AC component of the receptor potential to higher frequencies. Regarding the jitter of synaptic transmission, perhaps the anatomical convergence of many hair cell AZs onto one ANF acts to increase precision already in the auditory nerve, similar to the way that precision is enhanced in bushy cells and octopus cells of the mammalian cochlear nucleus through convergence of several ANFs. Afferent fibres with more presynaptic inputs may also have greater spontaneous spike rates, which may influence the upper frequency limits of phase locking beyond the lowpass filtering of the hair cell receptor potential. In a model of phase locking at IHC–ANF synapses of the cat, the f_c estimated for low spontaneous rate ANFs is lower than for high spontaneous rate ANFs, implying a synapse-specific contribution to lowpass filtering (Peterson & Heil, 2020).

Response phase and dependence on stimulus level. Finally, we consider the effect of stimulus level on response timing, or phase, for on-going sounds. How do interaural phase differences depend on interaural level differences? The mean response phase can be estimated from the position of the peak in the period histogram in response to a tone *in vivo* (Fig. 5*C*). In high spontaneous rate ANFs, the mean spike rate may change relatively little as SPL is increased, while the period histogram is modulated approximately sinusoidally with increasing vector strength. Remarkably, for low spontaneous rate ANFs, as SPL is increased by >90 dB the period histogram may remain modulated approximately sinusoidally as the mean spike rate increases by >10-fold (Rose *et al*. 1967). This phenomenon may be observed also in period histograms of EPSP timing as the amplitude of the sinusoidal hair cell receptor potential is increased from 5 mV to 20 mV in paired recordings of synaptic transfer in the amphibian papilla *ex vivo* (see Fig. 5 in Li *et al*. 2014).

How does the phase of phase-locking depend on the amplitude of a tone? The mean response phase will lead or lag with changing sound level in a systematic way that depends on the stimulus frequency relative to the 'null' frequency (Palmer & Shackleton, 2009). At the null frequency, in response to pure tones of increasing level, the mean spike rate can change from spontaneous to maximum while maintaining a relatively constant mean response phase (Anderson *et al*. 1971; Rose *et al*. 1967; Fuchs, 2005; Fig. 5*C*). The null frequency may be at CF, but often is not (Köppl, 1997). Below the null frequency, phase lags with increasing sound level. Above the null frequency, phase progressively leads with increasing sound level. These phase changes are thought to result largely from cochlear mechanics, as they can be seen at the level of basilar membrane motion and in the IHC receptor potential (Cheatham & Dallos, 2001; Robles & Ruggero, 2001). Remarkably, this same pattern is seen in birds (Gleich & Narins, 1988), and even in frogs, which have no basilar membrane but do have a tectorial membrane (Hillery & Narins, 1987), suggesting a mechanical vibratory phenomenon that is consistent across species. These phase changes may be predicted from the changes in cochlear filter shapes that occur with increasing sound level, as tuning of cochlear vibrations becomes broader (Tan & Carney, 2003). Some models of level encoding propose that the brain uses the pattern of phase variation across a population of ANFs to extract information outside of the limited dynamic range for changes in spike rate (Carney, 1994). Could the auditory system utilize these temporal features of sound representation across the array of ANFs to extend the range of information available in spike trains? If the phase shift in ANF discharge mirrors the phase shift in the IHC receptor potential due to phase shift in basilar membrane motion, *then the timing of synaptic transmission and spike generation should not introduce significant additional phase shifts with changes in SPL*. Here, we consider how this might be possible, given that the latency of synaptic transmission does depend on stimulus level and stimulus history based on general principles of conventional synapses (Felmy *et al*. 2003).

Synaptic latency at the guinea pig IHC–ANF synapse *in vivo* has been estimated to be ∼0.8 ms (Palmer & Russell, 1986). Similarly, paired recordings at room temperature in bullfrog have measured synaptic delays of about 0.5 ms from the peak of the Ca^{2+} tail current generated during a sinewave stimulus to the start of the EPSC (Li *et al*. 2014). As tone level increases and the hair cell becomes more depolarized, the accompanying faster increase in Ca^{2+} influx at the ribbon synapse might be expected to always trigger synaptic transmission faster and produce a phase lead in addition to any phase-shifts due to mechanics. The phase lead reflects a decrease in spike latency relative to some point in the excitatory phase of the periodic stimulus. How might phase leads be avoided as more Ca^{2+} channels are activated per stimulus cycle? If the speed of transmission is limited by the process of vesicle fusion with the plasma membrane after Ca^{2+} channels open, no matter how many channels open, then stimulus intensity may change the number of activated channels (and perhaps the strength of transmission) while having relatively little effect on the kinetics of exocytosis

and leaving phase unchanged (Moser *et al.* 2006). The expected result in this scenario may depend on the organization of readily releasable synaptic vesicles relative to voltage gated Ca^{2+} channels. If exocytosis is evoked by a nanometre-spaced Ca^{2+} channel (Brandt *et al.* 2005, Graydon *et al.* 2011), then a high-micromolar [Ca^{2+}] directly around the vesicle could make Ca^{2+} binding to the sensor occur at saturated rate regardless of the instantaneous receptor potential. Under these conditions, increases in sound level might increase release probability equally at all phases of the stimulus cycle such that mean preferred phase is maintained.

Finally, level-independent synaptic latency could arise from equal and opposite shifts in temporal properties of multiple underlying mechanisms. Paired recordings from IHC–ANF synapses suggested how multiple presynaptic mechanisms may combine to produce consistent release latencies across stimulus levels in response to the on-going part of periodic stimuli (Fig. 4). An increase in stimulus level will increase the presynaptic Ca^{2+} current driving exocytosis, tending to make transmission bigger and faster (Fig. 4*B, C* and *E*). At the same time, during that on-going stimulation, the increase in exocytosis depletes the pool of vesicles available for release, tending to make transmission smaller and slower (Fig. 4*A*, *B* and *D*; Goutman & Glowatzki, 2011). Thus, a balance between Ca^{2+}-dependent facilitation and vesicle supply-dependent depression may underlie the near phase constancy of release as an on-going periodic stimulus to the IHC is changed in intensity (Fig. 5*B*; Goutman, 2012). Alternatively, post-synaptic phase shifts of spike generation may offset presynaptic phase shifts of exocytosis. EPSC latency depends on depolarization level (Fig. 4*B*, *C* and *E*), with increasing level evoking faster transmission (see also Li *et al.* 2009). If EPSCs occur earlier but become smaller during high rates of release for loud sounds, then spike latency could remain relatively unshifted because smaller EPSCs evoke slower spike generation (Fig. 3*D*; Rutherford *et al.* 2012). Likewise, if EPSCs occurring later in the cycle tend to be larger at lower rates of release for low-level sounds, then again spike latency could remain relatively unshifted as larger EPSCs evoke faster spike generation. Interestingly, in the goldfish, a negative correlation was observed in the amplitudes of consecutive EPSPs in response to tones (Furukawa *et al.* 1978). During an on-going tone of constant level, if smaller EPSPs tend to be advanced in phase, and if larger EPSPs tend to be delayed in phase, then the spike generating mechanism may offset these phase shifts to achieve phase constancy at the level of the nerve discharge.

Conclusions and issues that remain unresolved

(1) The cochlea performs a Fourier transform on complex sounds in which vibrations of different frequencies and levels go to different places, to excite hair cells at different levels to evoke proportional spike rates in afferent fibres of the auditory nerve. In addition to this place-rate code for sound frequency and level, there exists a temporal code in which precise spike timing carries information. Cellular mechanisms determine the upper frequency limits for encoding of temporal information in sound. Temperature affects the encoding of sound and further study is required to understand how the kinetics of cellular mechanisms are influenced by temperature.

(2) The IHC–ANF ribbon synapse and spike generator are tightly coupled such that the timing of afferent spikes is determined by the timing of synaptic transmission. Spikes may be evoked by individual release events controlled by opening of just one or very few $Ca_V1.3$ channels in a 'nanodomain' coupling scenario that enhances temporal encoding for phase locking to low frequency sounds in the cochlear apex, where IHCs may be specialized to follow periodic stimulation with 'AC' receptor potentials. IHCs in the cochlear base may operate in a 'microdomain' coupling scenario that enhances encoding of sound level for high frequency sounds that exceed the speed limits of phase locking. In addition to direct promotion of microdomain coupling at the active zone, lower levels of cytoplasmic Ca^{2+} buffering in high-frequency IHCs may enhance encoding of sound level with summating (i.e. DC) potentials by reducing MET open probability which reduces the silent current, promotes asymmetry of transduction, hyperpolarizes the membrane potential, and lengthens the membrane time constant. Further study is required to understand the conditions under which nanodomain and microdomain coupling exist, including species differences and tonotopic variation.

(3) IHCs of the mammalian cochlea are not intrinsically frequency tuned like the electrically tuned sensory hair cells of non-mammalian vertebrates. Cochlear OHCs enhance sensitivity and frequency tuning through active processes in the bundle and basolateral membrane. High cytoplasmic Ca^{2+} buffering capacity increases MET open probability, increasing the silent current and depolarizing the membrane potential to promote symmetry of transduction and faster membrane time constants. This may eliminate a summating potential and allow OHCs to drive prestin-based electromotility on a cycle-by-cycle basis up to 100 kHz *in vivo*. The presence of tonotopic variation in OHC frequency tuning, including possible band-pass filtering by mechanisms in the hair bundle, requires further study.

(4) The upper frequency limit for phase locking in the auditory nerve is determined by the IHC membrane time constant and potentially the temporal jitter in synaptic transmission, as well. However, direct observations of temporal jitter in synaptic transmission have been made at room temperature and the jitter of transmission at mammalian body temperature, when the presynaptic $Ca_V1.3$ current is larger and faster, is not yet clear.

(5) There are important issues that remain to be resolved in the study of receptor potentials, synaptic transmission and afferent discharge. Many of them can be related to the use of reduced *ex vivo* preparations to infer mechanisms *in vivo* where there are differences in environmental conditions including ionic concentrations of fluids, maturity and temperature. Intracellular recordings from mature animals *in vivo* have provided ground truth data for aspects of signal processing but such experiments limit detailed study of cellular mechanisms due to relative inaccessibility of the inner ear in its native environment.

(6) The organ of Corti is a tight epithelium dividing two fluid compartments: basolateral membranes are exposed to perilymph containing high Ca^{2+} and low K^+; hair bundles are exposed to endolymph containing low Ca^{2+}, high K^+ and a 90 mV endolymphatic potential. Experiments on cellular mechanisms in explanted tissue are typically done with perilymph-like solutions, preventing study of the full system in an accessible way.

(7) To prolong the viability of the preparation, measurements are mostly made at room temperature. Non-linear effects of temperature on the mechanisms of presynaptic Ca^{2+} influx and efflux, synaptic transmission and postsynaptic spike generation may change temporal dynamics in unpredictable ways. Therefore, conclusions drawn from experiments at room temperature should be taken with caution and future experiments should assess mechanisms at body temperature.

(8) To avoid damage incurred by dissecting tissue from the temporal bone, the hardest bone in the body, younger animals are often preferred, particularly for the study of the fragile process of mechano-transduction. Therefore, conclusions drawn from experiments in young mammals should be taken with caution and future experiments should assess mechanisms in adults. Cold blooded species studied at room temperature provide valuable data in which temperature mismatch may be of less concern but generalizations to mammals should be made with caution.

(9) As IHC stimulation increases with sound level, Ca^{2+}-dependent synaptic facilitation and vesicle supply-dependent depression may offset to result in relative phase-invariance of ANF spiking relative to the phase of basilar membrane motion. This may be important for temporal encoding, but it is unclear exactly how the brain uses phase information in auditory signals.

(10) Heterogeneity in afferent signalling is introduced as the information in one IHC receptor potential diverges into the spike trains of several ANFs with different sound response properties. It is unclear if these heterogeneities combine randomly or in a pattern to increase diversity of encoding.

References

Anderson DJ, Rose JE, Hind JE & Brugge JF (1971). Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: frequency and intensity effects. *J Acoust Soc Am* **49**, 1131–1139.

Ashida G & Carr CE (2011). Sound localization: Jeffress and beyond. *Curr Opin Neurobiol* **21**, 745–751.

Ashmore JF (1987). A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J Physiol* **388**, 323–347.

Ashmore J (2010). The afferent synapse. In *The Oxford Handbook of Auditory Science: The Ear*, ed. Fuchs PA. Oxford University Press, Oxford, pp. 259–282.

Beurg M, Nam JH, Chen Q & Fettiplace R (2010). Calcium balance and mechanotransduction in rat cochlear hair cells. *J Neurophysiol* **104**, 18–34.

Beutner D & Moser T (2001). The presynaptic function of mouse cochlear inner hair cells during development of hearing. *J Neurosci* **21**, 4593–4599.

Boulet J & Bruce IC (2017). Predictions of the contribution of HCN half-maximal activation potential heterogeneity to variability in intrinsic adaptation of spiral ganglion neurons. *J Assoc Res Otolaryngol* **18**, 301–322.

Brandt A, Khimich D & Moser T (2005). Few $Ca_v1.3$ channels regulate the exocytosis of a synaptic vesicle at the hair cell ribbon synapse. *J Neurosci* **25**, 11577–11585.

Brandt A, Striessnig J & Moser T (2003). $Ca_v1.3$ channels are essential for development and presynaptic activity of cochlear inner hair cells. *J Neurosci* **23**, 10832–10840.

Brawer JR, Morest DK & Kane EC (1974). The neuronal architecture of the cochlear nucleus of the cat. *J Comp Neurol* **155**, 251–299.

Brown MC, Smith DI & Nuttall AL (1983). The temperature dependency of neural and hair cell responses evoked by high frequencies. *J Acoust Soc Am* **73**, 1662–1670.

Browne L, Smith KE & Jagger DJ (2017). Identification of persistent and resurgent sodium currents in spiral ganglion neurons cultured from the mouse cochlea. *eNeuro* **4**, ENEURO.0303-17.2017.

Brownell WE, Bader CR, Bertrand D & De Ribaupierre Y (1985). Evoked mechanical responses of isolated cochlear outer hair cells. *Science* **227**, 194–196.

Bulankina AV & Moser T (2012). Neural circuit development in the mammalian cochlea. *Physiology* **27**, 100–112.

Buran BN, Strenzke N, Neef A, Gundelfinger ED, Moser T & Liberman MC (2010). Onset coding is degraded in auditory nerve fibers from mutant mice lacking synaptic ribbons. *J Neurosci* **30**, 7587–7597.

Cao XJ & Oertel D (2010). Auditory nerve fibers excite targets through synapses that vary in convergence, strength, and shortterm plasticity. *J Neurophysiol* **104**, 2308–2320.

Carney LH (1994). Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness. *Hear Res* **76**, 31–44.

Carney LH (2018). Supra-threshold hearing and fluctuation profiles: implications for sensorineural and hidden hearing loss. *J Assoc Res Otolaryngol* **19**, 331–352.

Castellano- Muñoz M & Ricci AJ (2014). Role of intracellular calcium stores in hair-cell ribbon synapse. *Front Cell Neurosci* **8**, 162.

Cheatham MA & Dallos P (2001). Inner hair cell response patterns: implications for low-frequency hearing. *J Acoust Soc Am* **110**, 2034–2044.

Chen F, Zha D, Fridberger A, Zheng J, Choudhury N, Jacques SL, Wang RK, Shi X & Nuttall AL (2011). A differentially amplified motion in the ear for near-threshold sound detection. *Nat Neurosci* **14**, 770–774.

Chen M & von Gersdorff H (2019). How to build a fast and highly sensitive sound detector that remains robust to temperature shifts. *J Neurosci* **39**, 7260–7276.

Cho S, Li G-L & von Gersdorff H (2011). Recovery from short-term depression and facilitation is ultrafast and Ca^{2+} dependent at auditory hair cell synapses. *J Neurosci* **31**, 5682–5692.

Christensen-Dalsgaard J, Jørgensen MB & Kanneworff M (1998). Basic response characteristics of auditory nerve fibers in the grassfrog (*Rana temporaria*). *Hear Res* **119**, 155–163.

Christensen-Dalsgaard J & Narins PM (1993). Sound and vibration sensitivity of VIIIth nerve fibers in the frogs *Leptodactylus albilabris* and *Rana pipiens pipiens*. *J Comp Physiol A* **172**, 653–662.

Corey DP & Hudspeth AJ (1983). Kinetics of the receptor current in bullfrog saccular hair cells. *J Neurosci* **3**, 962–976.

Crawford AC & Fettiplace R (1980). The frequency selectivity of auditory nerve fibres and hair cells in the cochlea of the turtle. *J Physiol* **306**, 79–125.

Dallos P (1985). Response characteristics of mammalian cochlear hair cells. *J Neurosci* **5**, 1591–1608.

Dallos P (2008). Cochlear amplification, outer hair cells and prestin. *Curr Opin Neurobiol* **18**, 370–376.

Dallos P & Harris D (1978). Properties of auditory nerve responses in absence of outer hair cells. *J Neurophysiol* **41**, 365–383.

Davis H (1965). A model for transducer action in the cochlea. *Cold Spring Harb Symp Quant Biol* **30**, 181–190.

Delgutte B (1980). Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers. *J Acoust Soc Am* **68**, 843–857.

Dierich M, Altoè A, Koppelmann J, Evers S, Renigunta V, Schäfer MK, Naumann R, Verhulst S, Oliver D & Leitner MG (2020). Optimized tuning of auditory inner hair cells to encode complex sound through synergistic activity of six independent K^+ current entities. *Cell Rep* **32**, 107869.

Doll JC, Peng AW, Ricci AJ & Pruitt BL (2012). Faster than the speed of hearing: nanomechanical force probes enable the electromechanical observation of cochlear hair cells. *Nano Lett* **12**, 6107–6111.

Dulon D, Sugasawa M, Blanchet C & Erostegui C (1995). Direct measurements of Ca^{2+}-activated K^+ currents in inner hair cells of the guinea-pig cochlea using photolabile Ca^{2+} chelators. *Pflugers Arch* **430**, 365–373.

Eatock RA & Songer JE (2011). Vestibular hair cells and afferents: two channels for head motion signals. *Annu Rev Neurosci* **34**, 501–534.

Eggermont JJ (2001). Between sound and perception: reviewing the search for a neural code. *Hear Res* **157**, 1–42.

Engel J, Michna M, Platzer J & Striessnig J (2002). Calcium channels in mouse hair cells: function, properties and pharmacology. *Adv Otorhinolaryngol* **59**, 35–41.

Farris HE, Wells GB & Ricci AJ (2006). Steady-state adaptation of mechanotransduction modulates the resting potential of auditory hair cells, providing an assay for endolymph $[Ca^{2+}]$. *J Neurosci* **26**, 12526–12536.

Felmy F, Neher E & Schneggenburger R (2003). Probing the intracellular calcium sensitivity of transmitter release during synaptic facilitation. *Neuron* **37**, 801–811.

Fettiplace R (2017). Hair cell transduction, tuning, and synaptic transmission in the mammalian cochlea. *Compr Physiol* **7**, 1197–1227.

Fettiplace R & Fuchs PA (1999). Mechanisms of hair cell tuning. *Annu Rev Physiol* **61**, 809–834.

Fettiplace R & Kim KX (2014). The physiology of mechano-electrical transduction channels in hearing. *Physiol Rev* **94**, 951–986.

Fischer FP, Köppl C & Manley GA (1988). The basilar papilla of the barn owl *Tyto alba*: A quantitative morphological SEM analysis. *Hear Res* **34**, 87–101.

Frank G, Hemmert W & Gummer AW (1999). Limiting dynamics of high-frequency electromechanical transduction of outer hair cells. *Proc Natl Acad Sci U S A* **96**, 4420–4425.

Frank T, Khimich D, Neef A & Moser T (2009). Mechanisms contributing to synaptic Ca2+ signals and their heterogeneity in hair cells. *Proc Natl Acad Sci U S A* **106**, 4483–4488.

Frank T, Rutherford MA, Strenzke N, Neef A, Pangršič T, Khimich D, Fejtova A, Gundelfinger ED, Liberman MC, Harke B & Bryan KE (2010). Bassoon and the synaptic ribbon organize Ca^{2+} channels and vesicles to add release sites and promote refilling. *Neuron* **68**, 724–738.

Fuchs PA (2005). Time and intensity coding at the hair cell's ribbon synapse: The hair cell's afferent synapse. *J Physiol* **566**, 7–12.

Furukawa T, Hayashida Y & Matsuura S (1978). Quantal analysis of the size of excitatory post-synaptic potentials at synapses between hair cells and afferent nerve fibres in goldfish. *J Physiol* **276**, 211–226.

Furukawa T, Kuno M & Matsuura S (1982). Quantal analysis of a decremental response at hair cell-afferent fibre synapses in the goldfish sacculus. *J Physiol* **322**, 181–195.

Furukawa T & Matsuura S (1978). Adaptive rundown of excitatory post-synaptic potentials at synapses between hair cells and eight nerve fibres in the goldfish. *J Physiol* **276**, 193–209.

Gai Y, Doiron B, Kotak V & Rinzel J (2009). Noise-gated encoding of slow inputs by auditory brain stem neurons with a low-threshold K^+ current. *J Neurophysiol* **102**, 3447–3460.

Galambos R & Davis H (1943). The response of single auditory-nerve fibers to acoustic stimulation. *J Neurophysiol* **6**, 39–57.

Gebhart M, Juhasz-Vedres G, Zuccotti A, Brandt N, Engel J, Trockenbacher A, Kaur G, Obermair GJ, Knipper M, Koschak A & Striessnig J (2010). Modulation of $Ca_v1.3$ Ca^{2+} channel gating by Rab3 interacting molecule. *Mol Cell Neurosci* **44**, 246–259.

Gleich O & Narins PM (1988). The phase response of primary auditory afferents in a songbird (*Sturnus vulgaris* L.). *Hear Res* **32**, 81–91.

Glowatzki E & Fuchs PA (2002). Transmitter release at the hair cell ribbon synapse. *Nat Neurosci* **5**, 147–154.

Goldberg JM & Brown PB (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* **32**, 613–636.

Golding, NL, Robertson D & Oertel D (1995). Recordings from slices indicate that octopus cells of the cochlear nucleus detect coincident firing of auditory nerve fibers with temporal precision. *J Neurosci* **15**, 3138–3153.

Goutman JD (2012). Transmitter release from cochlear hair cells is phase locked to cyclic stimuli of different intensities and frequencies. *J Neurosci* **32**, 17025–17036.

Goutman JD & Glowatzki E (2007). Time course and calcium dependence of transmitter release at a single ribbon synapse. *Proc Natl Acad Sci U S A* **104**, 16341–16346.

Goutman JD & Glowatzki E (2011). Short-term facilitation modulates size and timing of the synaptic response at the inner hair cell ribbon synapse. *J Neurosci* **31**, 7974–7981.

Grant L & Fuchs P (2008). Calcium- and calmodulin-dependent inactivation of calcium channels in inner hair cells of the rat cochlea. *J Neurophysiol* **99**, 2183–2193.

Grant L, Yi E & Glowatzki E (2010). Two modes of release shape the postsynaptic response at the inner hair cell ribbon synapse. *J Neurosci* **30**, 4210–4220.

Graydon CW, Cho S, Diamond JS, Kachar B, von Gersdorff H & Grimes WN (2014). Specialized postsynaptic morphology enhances neurotransmitter dilution and high-frequency signaling at an auditory synapse. *J Neurosci* **34**, 8358–8372.

Graydon CW, Cho S, Li G-L, Kachar B & von Gersdorff H (2011). Sharp Ca^{2+} nanodomains beneath the ribbon promote highly synchronous multivesicular release at hair cell synapses. *J Neurosci* **31**, 16637–16650.

Greenwood DD (1990). A cochlear frequency-position function for several species—29 years later. *J Acoust Soc Am* **87**, 2592–2605.

Greenwood DD & Joris PX (1996). Mechanical and "temporal" filtering as codeterminants of the response by cat primary fibers to amplitude-modulated signals. *J Acoust Soc Am* **99**, 1029–1039.

Grothe B, Pecka M & McAlpine D (2010). Mechanisms of sound localization in mammals. *Physiol Rev* **90**, 983–1012.

Guinan JJ Jr (2010). Cochlear efferent innervation and function. *Curr Opin Otolaryngol Head Neck Surg* **18**, 447–453.

Guo W, Chambers AR, Darrow KN, Hancock KE, Shinn-Cunningham BG & Polley DB (2012). Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types. *J Neurosci* **32**, 9159–9172.

Hackney CM, Mahendrasingam S, Penn A & Fettiplace R (2005). The concentrations of calcium buffering proteins in mammalian cochlear hair cells. *J Neurosci* **25**, 7867–7875.

Hardie J & Lee A (2016). Decalmodulation of Cav1 channels by CaBPs. *Channels* **10**, 33–37.

Heil P & Irvine DRF (1997). First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J Neurophysiol* **78**, 2438–2454.

Heil P & Neubauer H (2001). Temporal integration of sound pressure determines thresholds of auditory-nerve fibers. *J Neurosci* **21**, 7404–7415.

Heil P & Neubauer H (2003). A unifying basis of auditory thresholds based on temporal summation. *Proc Natl Acad Sci U S A* **100**, 6151–6156.

Heil P & Peterson AJ (2015). Basic response properties of auditory nerve fibers: a review. *Cell Tissue Res* **361**, 129–158.

Heil P & Peterson AJ (2017). Spike timing in auditory-nerve fibers during spontaneous activity and phase locking: ANF spontaneous activity and phase locking. *Synapse* **71**, 5–36.

Hillery CM & Narins PM (1987). Frequency and time domain comparison of low-frequency auditory fiber responses in two anuran amphibians. *Hear Res* **25**, 233–248.

Hodgkin AL (1948). The local electric changes associated with repetitive action in a non-medullated axon. *J Physiol* **107**, 165–181.

Hossain WA (2005). Where is the spike generator of the cochlear nerve? Voltage-gated sodium channels in the mouse cochlea. *J Neurosci* **25**, 6857–6868.

Housley GD & Ashmore JF (1992). Ionic currents of outer hair cells isolated from the guinea-pig cochlea. *J Physiol* **448**, 73–98.

Hudspeth AJ (2014). Integrating the active process of hair cells with cochlear function. *Nat Rev Neurosci* **15**, 600–614.

Huet A, Desmadryl G, Justal T, Nouvian R, Puel J-L & Bourien J (2018). The interplay between spike-time and spike-rate modes in the auditory nerve encodes tone-in-noise threshold. *J Neurosci* **38**, 5727–5738.

Huguet G, Meng X & Rinzel J (2017). Phasic firing and coincidence detection by subthreshold negative feedback: Divisive or subtractive or, better, both. *Front Comput Neurosci* **11**, 3.

Izhikevich EM (2007). *Dynamical Systems in Neuroscience*. MIT press, Cambridge, MA.

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

Johnson SL (2015). Membrane properties specialize mammalian inner hair cells for frequency or intensity encoding. *Elife* **4**, e08177.

Johnson SL, Beurg M, Marcotti W & Fettiplace R (2011). Prestin-driven cochlear amplification is not limited by the outer hair cell membrane time constant. *Neuron* **70**, 1143–1154.

Johnson SL, Forge A, Knipper M, Münkner S & Marcotti W (2008). Tonotopic variation in the calcium dependence of neurotransmitter release and vesicle pool replenishment at mammalian auditory ribbon synapses. *J Neurosci* **28**, 7670–7678.

Johnson SL, Franz C, Knipper M & Marcotti W (2009). Functional maturation of the exocytotic machinery at gerbil hair cell ribbon synapses. *J Physiol* **587**, 1715–1726.

Johnson SL & Marcotti W (2008). Biophysical properties of $Ca_V1.3$ calcium channels in gerbil inner hair cells. *J Physiol* **586**, 1029–1042.

Johnson SL, Marcotti W & Kros CJ (2005). Increase in efficiency and reduction in Ca^{2+} dependence of exocytosis during development of mouse inner hair cells. *J Physiol* **563**, 177–191.

Johnson SL, Olt J, Cho S, Von Gersdorff H & Marcotti W (2017). The coupling between Ca^{2+} channels and the exocytotic Ca^{2+} sensor at hair cell ribbon synapses varies tonotopically along the mature cochlea. *J Neurosci* **37**, 2471–2484.

Johnson SL, Safieddine S, Mustapha M & Marcotti W (2019). Hair cell afferent synapses: function and dysfunction. *Cold Spring Harb Perspect Med* **9**, a033175.

Joris PX, Carney LH, Smith PH & Yin TC (1994*a*). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J Neurophysiol* **71**, 1022–1036.

Joris PX, Schreiner CE & Rees A (2004). Neural processing of amplitude-modulated sounds. *Physiol Rev* **84**, 541–577.

Joris PX, Smith PH & Yin TC (1994*b*). Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *J Neurophysiol* **71**, 1037–1051.

Joris PX & Verschooten E (2013). On the limit of neural phase locking to fine structure in humans. *Adv Exp Med Biol* **787**, 101–108.

Joris PX & Yin TC (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am* **91**, 215–232.

Jung S, Oshima-Takago T, Chakrabarti R, Wong AB, Jing Z, Yamanbaeva G, Picher MM, Wojcik SM, Göttfert F, Predoehl F & Michel K (2015). Rab3-interacting molecules 2α and 2β promote the abundance of voltage-gated $Ca_V1.3$ Ca^{2+} channels at hair cell active zones. *Proc Natl Acad Sci U S A* **112**, E3141–E3149.

Keen EC & Hudspeth AJ (2006). Transfer characteristics of the hair cell's afferent synapse. *Proc Natl Acad Sci U S A* **103**, 5537–5542.

Kennedy HJ, Crawford AC & Fettiplace R (2005). Force generation by mammalian hair bundles supports a role in cochlear amplification. *Nature* **433**, 880–883.

Kiang NYS, Watanabe T, Thomas EC & Clark LF (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. Research Monograph No. 35. MIT Press, Cambridge, MA.

Kiang NYS(1984) Peripheral neural processing of auditory information. In *Handbook of Physiology*, Section 1, The Nervous System, Vol. III, Sensory Processes, ed. Brookhart JM & Mountcastle V, pp. 639–674.American Physiological Society, Bethesda.

Kim DO & Molnar CE (1979). A population study of cochlear nerve fibers: comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones. *J Neurophysiol* **42**, 16–30.

Kim JH, Kushmerick C & von Gersdorff H (2010). Presynaptic resurgent Na^+ currents sculpt the action potential waveform and increase firing reliability at a CNS nerve terminal. *J Neurosci* **30**, 15479–15490.

Kim KX & Rutherford MA (2016). Maturation of Na_V and K_V channel topographies in the auditory nerve spike initiator before and after developmental onset of hearing function. *J Neurosci* **36**, 2111–2118.

Kimberley BP, Brown DK & Eggermont JJ (1993). Measuring human cochlear traveling wave delay using distortion product emission phase responses. *J Acoust Soc Am* **94**, 1343–1350.

Klumpp RG & Eady HR (1956). Some measurements of interaural time difference thresholds. *J Acoust Soc Am* **28**, 859–860.

Knudsen EI & Konishi M (1979). Mechanisms of sound localization in the barn owl (*Tyto alba*). *J Comp Physiol* **133**, 13–21.

Köppl C (1997). Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, *Tyto alba*. *J Neurosci* **17**, 3312–3321.

Köppl C (2011). Birds – same thing, but different? Convergent evolution in the avian and mammalian auditory systems provides informative comparative models. *Hearing Res* **273**, 65–71.

Köppl C & Manley GA (2019). A functional perspective on the evolution of the cochlea. *Cold Spring Harb Perspect Med* **9**, a033241.

Krinner S, Butola T, Jung S, Wichmann C & Moser T (2017). RIM-binding protein 2 promotes a large number of $Ca_V1.3$ Ca^{2+}-channels and contributes to fast synaptic vesicle replenishment at hair cell active zones. *Front Cell Neurosci* **11**, 334.

Kroll J, Jaime Tobón LM, Vogl C, Neef J, Kondratiuk I, König M, Strenzke N, Wichmann C, Milosevic I & Moser T (2019). Endophilin-A regulates presynaptic Ca^{2+} influx and synaptic vesicle recycling in auditory hair cells. *EMBO J* **38**, e100116.

Kros C & Crawford A (1990). Potassium currents in inner hair cells isolated from the guinea-pig cochlea. *J Physiol* **421**, 263–291.

Kros CJ, Ruppersberg JP & Rüsch A (1998). Expression of a potassium current in inner hair cells during development of hearing in mice. *Nature* **394**, 281–284.

Kujawa SG & Liberman MC (2009). Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. *J Neurosci* **29**, 14077–14085.

Langner G (1992). Periodicity coding in the auditory system. *Hear Res* **60**, 115–142.

Langner G, Sams M, Heil P & Schulze H (1997). Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. *J Comp Physiol A* **181**, 665–676.

Levic S, Bouleau Y & Dulon D (2011). Developmental acquisition of a rapid calcium-regulated vesicle supply allows sustained high rates of exocytosis in auditory hair cells. *PLoS One* **6**, e25714.

Li G-L, Cho S & von Gersdorff H (2014). Phase-locking precision is enhanced by multiquantal release at an auditory hair cell ribbon synapse. *Neuron* **83**, 1404–1417.

Li G-L, Keen E, Andor-Ardo D, Hudspeth AJ & von Gersdorff H (2009). The unitary event underlying multiquantal EPSCs at a hair cell's ribbon synapse. *J Neurosci* **29**, 7558–7568.

Liberman MC (1980). Morphological differences among radial afferent fibers in the cat cochlea: an electron-microscopic study of serial sections. *Hear Res* **3**, 45–63.

Liberman MC (1982). Single-neuron labeling in the cat auditory nerve. *Science* **216**, 1239–1241.

Liberman MC (1991). Central projections of auditory-nerve fibers of differing spontaneous rate. I. Anteroventral cochlear nucleus. *J Comp Neurol* **313**, 240–258.

Liberman MC (1993). Central projections of auditory nerve fibers of differing spontaneous rate, II: Posteroventral and dorsal cochlear nuclei. *J Comp Neurol* **327**, 17–36.

Liberman MC, Gao J, He DZ, Wu X, Jia S & Zuo J (2002). Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature* **419**, 300–304.

Licklider JCR (1954). "Periodicity" pitch and "place" pitch. *J Acoust Soc Am* **26**, 945–945.

Lorenzi C, Gilbert G, Carn H, Garnier S & Moore BCJ (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci U S A* **103**, 18866–18869.

Lorenzi C & Moore BC (2007). Role of temporal envelope and fine structure cues in speech perception: A review. In *Proceedings of the International Symposium on Auditory and Audiological Research*, vol. **1**, pp. 263–272.

Liu Q, Lee E & Davis RL (2014*a*). Heterogeneous intrinsic excitability of murine spiral ganglion neurons is determined by Kv1 and HCN channels. *Neuroscience* **257**, 96–110.

Liu Q, Manis PB & Davis RL (2014*b*). I_h and HCN channels in murine spiral ganglion neurons: tonotopic variation, local heterogeneity, and kinetic model. *J Assoc Res Otolaryngol* **15**, 585–599.

Lysakowski A, Gaboyard-Niay S, Calin-Jageman I, Chatlani S, Price SD & Eatock RA (2011). Molecular microdomains in a sensory terminal, the vestibular calyx ending. *J Neurosci* **31**, 10101–10114.

Magistretti J, Spaiardi P, Johnson SL & Masetto S (2015). Elementary properties of Ca^{2+} channels and their influence on multivesicular release and phase-locking at auditory hair cell ribbon synapses. *Front Cell Neurosci* **9**, 123.

Mammano F & Ashmore JF (1996). Differential expression of outer hair cell potassium currents in the isolated cochlea of the guinea-pig. *J Physiol* **496**, 639–646.

Marcotti W, Johnson SL, Holley MC & Kros CJ (2003*a*). Developmental changes in the expression of potassium currents of embryonic, neonatal and mature mouse inner hair cells. *J Physiol* **548**, 383–400.

Marcotti W, Johnson SL, Rüsch A & Kros CJ (2003*b*). Sodium and calcium currents shape action potentials in immature mouse inner hair cells. *J Physiol* **552**, 743–761.

Marcotti W & Kros CJ (1999). Developmental expression of the potassium current $I_{K,n}$ contributes to maturation of mouse outer hair cells. *J Physiol* **520**, 653–660.

McAlpine D, Jiang D & Palmer AR (2001). A neural code for low-frequency sound localization in mammals. *Nat Neurosci* **4**, 396–401.

McFadden D (1988). Failure of a missing-fundamental complex to interact with masked and unmasked pure tones at its fundamental frequency. *Hear Res* **32**, 23–39.

McGinley MJ & Oertel D (2006). Rate thresholds determine the precision of temporal integration in principal cells of the ventral cochlear nucleus. *Hear Res* **216–217**, 52–63.

Meyer AC & Moser T (2010). Structure and function of cochlear afferent innervation. *Curr Opin Otolaryngol Head Neck Surg* **18**, 441–446.

Mo Z-L, Adamson CL & Davis RL (2002). Dendrotoxin-sensitive K^+ currents contribute to accommodation in murine spiral ganglion neurons. *J Physiol* **542**, 763–778.

Moore BCJ (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Otolaryngol* **9**, 399–406.

Moser T & Beutner D (2000). Kinetics of exocytosis and endocytosis at the cochlear inner hair cell afferent synapse of the mouse. *Proc Natl Acad Sci U S A* **97**, 883–888.

Moser T, Neef A & Khimich D (2006). Mechanisms underlying the temporal precision of sound coding at the inner hair cell ribbon synapse: Temporal precision of synaptic sound coding. *J Physiol* **576**, 55–62.

Moser T & Strenzke N (2015). Synaptic encoding and processing of auditory information in physiology and disease. *Hear Res* **330**, 155–156.

Müller M (1996). The cochlear place-frequency map of the adult and developing Mongolian gerbil. *Hear Res* **94**, 148–156.

Müller M, von Hünerbein K, Hoidis S & Smolders JW (2005). A physiological place–frequency map of the cochlea in the CBA/J mouse. *Hear Res* **202**, 63–73.

Muniak MA, Connelly CJ, Suthakar K, Milinkeviciute G, Ayeni FE & Ryugo DK (2016). Central projections of spiral ganglion neurons. In *The Primary Auditory Neurons of the Mammalian Cochlea, Springer Handbook of Auditory Research*, ed. Dabdoub A, Fritzsch B, Popper AN & Fay RR. Springer, New York.

Narayan SS, Temchin AN, Recio A & Ruggero MA (1998). Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae. *Science* **282**, 1882–1884.

Neef J, Gehrt A, Bulankina AV, Meyer AC, Riedel D, Gregg RG, Strenzke N & Moser T (2009). The Ca^{2+} channel subunit $\beta 2$ regulates Ca^{2+} channel abundance and function in inner hair cells and is required for hearing. *J Neurosci* **29**, 10730–10740.

Neef J, Urban NT, Ohn T-L, Frank T, Jean P, Hell SW, Willig KI & Moser T (2018). Quantitative optical nanophysiology of Ca^{2+} signaling at inner hair cell active zones. *Nat Commun* **9**, 290.

Nie K, Barco A & Zeng FG (2006). Spectral and temporal cues in cochlear implant speech perception. *Ear Hear* **27**, 208–217.

Nouvian R (2007). Temperature enhances exocytosis efficiency at the mouse inner hair cell ribbon synapse: Temperature effects on hair cell synaptic function. *J Physiol* **584**, 535–542.

Nouvian R, Eybalin M & Puel JL (2015). Cochlear efferents in developing adult and pathological conditions. *Cell Tissue Res* **361**, 301–309.

Oertel D, Bal R, Gardner SM, Smith PH & Joris PX (2000). Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. *Proc Natl Acad Sci USA* **97**, 11773–11779.

Oertel D, Shatadal S & Cao XJ (2008). In the ventral cochlear nucleus $K_v1.1$ and subunits of HCN1 are colocalized at surfaces of neurons that have low-voltage-activated and hyperpolarization-activated conductances. *Neuroscience* **154**, 77–86.

Ohlemiller KK & Siegel JH (1992). The effects of moderate cooling on gross cochlear potentials in the gerbil: basal and apical differences. *Hear Res* **63**, 79–89.

Ohlemiller KK & Siegel JH (1994). Cochlear basal and apical differences reflected in the effects of cooling on responses of single auditory nerve fibers. *Hear Res* **80**, 174–190.

Ohlemiller KK & Siegel JH (1998). Temporal aspects of the effects of cooling on responses of single auditory nerve fibers. *Hear Res* **123**, 78–86.

Ohn T-L, Rutherford MA, Jing Z, Jung S, Duque-Afonso CJ, Hoch G, Picher MM, Scharinger A, Strenzke N & Moser T (2016). Hair cells use active zones with different voltage dependence of Ca^{2+} influx to decompose sounds into complementary neural codes. *Proc Natl Acad Sci U S A* **113**, E4716–E4725.

Ortner NJ, Pinggera A, Hofer NT, Siller A, Brandt N, Raffeiner A, Vilusic K, Lang I, Blum K, Obermair GJ & Stefan E (2020). RBP2 stabilizes slow $Ca_v1.3$ Ca^{2+} channel inactivation properties of cochlear inner hair cells. *Pflugers Arch* **472**, 3–25.

Oxenham AJ & Kreft HA (2014). Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing. *Trends Hearing* **18**, 233121651455378.

Palmer AR (1982). Encoding of rapid amplitude fluctuations by cochlear-nerve fibres in the guinea-pig. *Arch Otorhinolaryngol* **236**, 197–202.

Palmer AR, & Russell IJ (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear Res* **24**, 1–15.

Palmer AR & Shackleton TM (2009). Variation in the phase of response to low-frequency pure tones in the Guinea pig auditory nerve as functions of stimulus level and frequency. *J Assoc Res Otolaryngol* **10**, 233–250.

Pangršič T, Gabrielaitis M, Michanski S, Schwaller B, Wolf F, Strenzke N & Moser T (2015). EF-hand protein Ca^{2+} buffers regulate Ca^{2+} influx and exocytosis in sensory hair cells. *Proc Natl Acad Sci U S A* **112**, E1028–E1037.

Peterson AJ & Heil P (2018). A simple model of the inner-hair-cell ribbon synapse accounts for mammalian auditory-nerve-fiber spontaneous spike times. *Hear Res* **363**, 1–27.

Peterson AJ & Heil P (2020). Phase locking of auditory nerve fibers: the role of lowpass filtering by hair cells. *J Neurosci* **40**, 4700–4714.

Peterson AJ, Huet A, Bourien J, Puel JL & Heil P (2018). Recovery of auditory-nerve-fiber spike amplitude under natural excitation conditions. *Hear Res* **370**, 248–263.

Peterson AJ, Irvine DRF & Heil P (2014). A model of synaptic vesicle-pool depletion and replenishment can account for the interspike interval distributions and nonrenewal properties of spontaneous spike trains of auditory-nerve fibers. *J Neurosci* **34**, 15097–15109.

Petitpré C, Bourien J, Wu H, Diuba A, Puel JL & Lallemend F (2020). Genetic and functional diversity of primary auditory afferents. *Curr Opin Physiol* **18**, 85–94,

Petitpré C, Wu H, Sharma A, Tokarska A, Fontanet P, Wang Y, Helmbacher F, Yackle K, Silberberg G, Hadjab S & Lallemend F (2018). Neuronal heterogeneity and stereotyped connectivity in the auditory afferent system. *Nat Commun* **9**, 3691.

Picher MM, Gehrt A, Meese S, Ivanovic A, Predoehl F, Jung S, Schrauwen I, Dragonetti AG, Colombo R, Van Camp G, Strenzke N & Moser T (2017*a*). Ca^{2+} -binding protein 2 inhibits Ca^{2+}-channel inactivation in mouse inner hair cells. *Proc Natl Acad Sci U S A* **114**, E1717–E1726.

Picher MM, Oprişoreanu AM, Jung S, Michel K, Schoch S & Moser T (2017*b*). Rab interacting molecules 2 and 3 directly interact with the pore-forming $Ca_V1.3$ Ca^{2+} channel subunit and promote its membrane expression. *Front Cell Neurosci* **11**, 160.

Platzer J, Engel J, Schrott-Fischer A, Stephan K, Bova S, Chen H, Zheng H & Striessnig J (2000). Congenital deafness and sinoatrial node dysfunction in mice lacking class D L-type Ca2+ channels. *Cell* **102**, 89–97.

Portfors CV & von Gersdorff H (2013). Macrocircuits for sound localization use leaky coincidence detectors and specialized synapses. *Neuron* **78**, 755–757.

Prescott SA, De Koninck Y & Sejnowski TJ (2008). Biophysical basis for three distinct dynamical mechanisms of action potential initiation. *PLoS Comput Biol* **4**, e1000198.

Puel JL, Ruel J, d'Aldin CG & Pujol R (1998). Excitotoxicity and repair of cochlear synapses after noise-trauma induced hearing loss. *Neuroreport* **9**, 2109–2114.

Raman IM, Sprunger LK, Meisler MH & Bean BP (1997). Altered subthreshold sodium currents and disrupted firing patterns in purkinje neurons of *Scn8a* mutant mice. *Neuron* **19**, 881–891.

Reichenbach T & Hudspeth AJ (2014). The physics of hearing: fluid mechanics and the active process of the inner ear. *Rep Prog Phys* **77**, 076601.

Reijntjes DO, Lee JH, Park S, Schubert NM, van Tuinen M, Vijayakumar S, Jones TA, Jones SM, Gratton MA, Xia XM & Yamoah EN (2019). Sodium-activated potassium channels shape peripheral auditory function and activity of the primary auditory neurons in mice. *Sci Rep* **9**, 1–18.

Ricci AJ, Crawford AC & Fettiplace R (2003). Tonotopic variation in the conductance of the hair cell mechano-transducer channel. *Neuron* **40**, 983–990.

Ricci AJ, Kennedy HJ, Crawford AC & Fettiplace R (2005). The transduction channel filter in auditory hair cells. *J Neurosci* **25**, 7831–7839.

Ricci AJ, Wu YC & Fettiplace R (1998). The endogenous calcium buffer and the time course of transducer adaptation in auditory hair cells. *J Neurosci* **18**, 8261–8277.

Richardson GP, de Monvel JB & Petit C (2011). How the genetics of deafness illuminates auditory physiology. *Annu Rev Physiol* **73**, 311–334.

Robertson D & Paki B (2002). Role of L-type Ca^{2+} channels in transmitter release from mammalian inner hair cells. II. Single-neuron activity. *J Neurophysiol* **87**, 2734–2740.

Robles L & Ruggero MA (2001). Mechanics of the mammalian cochlea. *Physiol Rev* **81**, 1305–1352.

Rose JE, Brugge JF, Anderson DJ & Hind JE (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol* **30**, 769–793.

Rubio ME (2018). Microcircuits of the Ventral Cochlear Nucleus. In *The Mammalian Auditory Pathways: Synaptic Organization and Microcircuits*, ed. Oliver DL, Cant N, Fay RR & Popper AN. Springer Science and Business Media, New York.

Russell IJ, Drexl M, Foeller E, Vater M & Kössl M (2003). Synchronization of a nonlinear oscillator: processing the Cf component of the echo-response signal in the cochlea of the mustached bat. *J Neurosci* **23**, 9508–9518.

Russell IJ & Sellick PM (1978). Intracellular studies of hair cells in the mammalian cochlea. *J Physiol* **284**, 261–290.

Russell IJ & Sellick PM (1983). Low-frequency characteristics of intracellularly recorded receptor potentials in guinea-pig cochlear hair cells. *J Physiol* **338**, 179–206.

Rutherford MA (2015) Resolving the structure of the inner ear ribbon synapses with STED microscopy. *Synapse* **69**, 242–255.

Rutherford MA, Chapochnikov NM & Moser T (2012). Spike encoding of neurotransmitter release timing by spiral ganglion neurons of the cochlea. *J Neurosci* **32**, 4773–4789.

Rutherford MA & Pangršič T (2012). Molecular anatomy and physiology of exocytosis in sensory hair cells. *Cell Calcium* **52**, 327–337.

Rutherford & Moser (2016). The ribbon synapse between type I spiral ganglion neurons and inner hair cells. In: *The Primary Auditory Neurons of the Mammalian Cochlea. Springer Handbook of Auditory Research*, ed. Dabdoub A, Fritzsch B, Popper AN, Fay RR. Springer, New York.

Rutherford MA & Roberts WM (2006). Frequency selectivity of synaptic exocytosis in frog saccular hair cells. *Proc Natl Acad Sci U S A* **103**, 2898–2903.

Ryugo DK (2008). Projections of low spontaneous rate, high threshold auditory nerve fibers to the small cell cap of the cochlear nucleus in cats. *Neuroscience* **154**, 114–126.

Salminen NH, Tiitinen H, Yrttiaho S & May PJ (2010). The neural code for interaural time difference in human auditory cortex. *J Acoust Soc Am* **127**, EL60–EL65.

Santos-Sacchi J (1993). Voltage-dependent ionic conductances of type I spiral ganglion cells from the guinea pig inner ear. *J Neurosci* **13**, 3599–3611.

Schmidt RS & Fernández C (1962). Labyrinthine DC potentials in representative vertebrates. *J Cell Comp Physiol* **59**, 311–322.

Schnee ME, Lawton DM, Furness DN, Benke TA & Ricci AJ (2005). Auditory hair cell-afferent fiber synapses are specialized to operate at their best frequencies. *Neuron* **47**, 243–254.

Schnee ME, Castellano-Muñoz M & Ricci AJ (2013). Response properties from turtle auditory hair cell afferent fibers suggest spike generation is driven by synchronized release both between and within synapses. *J Neurophysiol* **110**, 204–220.

Schnee ME, Santos-Sacchi J, Castellano-Muñoz M, Kong JH & Ricci AJ (2011). Calcium-dependent synaptic vesicle trafficking underlies indefatigable release at the hair cell afferent fiber synapse. *Neuron* **70**, 326–338.

Schouten JF (1970). The residue revisited. In *Frequency Analysis and Periodicity Detection in Hearing*, ed. Plomp R & Smoorenburg GF, pp. 41–58.Sijthoff, Leiden.

Scott LL, Mathews PJ & Golding NL (2005). Posthearing developmental refinement of temporal processing in principal neurons of the medial superior olive. *J Neurosci* **25**, 7887–7895.

Sellick PM, Patuzzi R & Johnstone BM (1982). Modulation of responses of spiral ganglion cells in the guinea pig cochlea by low frequency sound. *Hear Res* **7**, 199–221.

Sewell WF (1984). The relation between the endocochlear potential and spontaneous activity in auditory nerve fibres of the cat. *J Physiol* **347**, 685–696.

Shamma S (2001). On the role of space and time in auditory processing. *Trends Cogn Sci* **5**, 340–348.

Shannon RV (2002). The relative importance of amplitude, temporal, and spectral cues for cochlear implant processor design. *Am J Audiol* **11**, 124–127,

Shrestha BR, Chia C, Wu L, Kujawa SG, Liberman MC & Goodrich LV (2018). Sensory neuron diversity in the inner ear is shaped by activity. *Cell* **174**, 1229–1246.e17.

Siegel JH (1992). Spontaneous synaptic potentials from afferent terminals in the guinea pig cochlea. *Hear Res* **59**, 85–92.

Simmons DD, Bertolotto C & Narins PM (1992). Innervation of the amphibian and basilar papillae in the leopard frog: reconstructions of single labeled fibers. *J Comp Neurol* **322**, 191–200.

Smith KE, Browne L, Selwood DL, McAlpine D & Jagger DJ (2015). Phosphoinositide modulation of heteromeric Kv1 channels adjusts output of spiral ganglion neurons from hearing mice. *J Neurosci* **35**, 11221–11232.

Songer JE & Eatock RA (2013). Tuning and timing in mammalian type I hair cells and calyceal synapses. *J Neurosci* **33**, 3706–3724.

Spoendlin H (1972). Innervation densities of the cochlea. *Acta Otolaryngol* **73**, 235–248.

Suga N (1965). Functional properties of auditory neurones in the cortex of echo-locating bats. *J Physiol* **181**, 671–700.

Sun S, Babola T, Pregernig G, So KS, Nguyen M, Su SSM, Palermo AT, Bergles DE, Burns JC & Müller U (2018). Hair cell mechanotransduction regulates spontaneous activity and spiral ganglion subtype specification in the auditory system. *Cell* **174**, 1247–1263.e15.

Taberner AM & Liberman MC (2005). Response properties of single auditory nerve fibers in the mouse. *J Neurophysiol* **93**, 557–569.

Tan Q & Carney LH (2003). A phenomenological model for the responses of auditory-nerve fibers. II. Nonlinear tuning with a frequency glide. *J Acoust Soc Am* **114**, 2007–2020.

Tertrais M, Bouleau Y, Emptoz A, Belleudy S, Sutton RB, Petit C, Safieddine S & Dulon D (2019). Viral transfer of mini-otoferlins partially restores the fast component of exocytosis and uncovers ultrafast endocytosis in auditory hair cells of otoferlin knock-out mice. *J Neurosci* **39**, 3394–3411.

Troconis EL, Ordoobadi AJ, Sommers TF, Aziz-Bose R, Carter AR & Trapani JG (2017). Intensity-dependent timing and precision of startle response latency in larval zebrafish: Intensity-dependent startle responses. *J Physiol* **595**, 265–282.

Tsuji J & Liberman MC (1997). Intracellular labeling of auditory nerve fibers in guinea pig: central and peripheral projections. *J Comp Neurol* **381**, 188–202.

Vavakou A, Cooper NP & van der Heijden M (2019). The frequency limit of outer hair cell motility measured in vivo. *Elife* **8**, e47667.

Ventura CM & Kalluri R (2019). Enhanced activation of HCN channels reduces excitability and spike-timing regularity in maturing vestibular afferent neurons. *J Neurosci* **39**, 2860–2876.

Von Békésy G & Wever EG (1960). *Experiments in Hearing*. McGraw-Hill, New York.

Wan G & Corfas G (2017). Transient auditory nerve demyelination as a new mechanism for hidden hearing loss. *Nat Commun* **8**, 14487.

Wang Y, Fallah E & Olson ES (2019). Adaptation of cochlear amplification to low endocochlear potential. *Biophys J* **116**, 1769–1786.

Wichmann C (2015). Molecularly and structurally distinct synapses mediate reliable encoding and processing of auditory information. *Hear Res* **330**, 178–190.

Wichmann C & Moser T (2015). Relating structure and function of inner hair cell ribbon synapses. *Cell Tissue Res* **361**, 95–114.

Wittig JH & Parsons TD (2008). Synaptic ribbon enables temporal precision of hair cell afferent synapse by increasing the number of readily releasable vesicles: a modeling study. *J Neurophysiol* **100**, 1724–1739.

Wong AB, Jing Z, Rutherford MA, Frank T, Strenzke N, & Moser T (2013). Concurrent maturation of inner hair cell synaptic Ca^{2+} influx and auditory nerve spontaneous activity around hearing onset in mice. *J Neurosci* **33**, 10661–10666.

Wong AB, Rutherford MA, Gabrielaitis M, Pangršič T, Göttfert F, Frank T, Michanski S, Hell S, Wolf F, Wichmann C & Moser T (2014). Developmental refinement of hair cell synapses tightens the coupling of Ca^{2+} influx to exocytosis. *EMBO J* **33**, 247–264.

Wu JS, Young ED & Glowatzki E (2016). Maturation of spontaneous firing properties after hearing onset in rat auditory nerve fibers: spontaneous rates, refractoriness, and interfiber correlations. *J Neurosci* **36**, 10584–10597.

Wu YC, Ricci AJ & Fettiplace R (1999). Two components of transducer adaptation in auditory hair cells. *J Neurophysiol* **82**, 2171–2181.

Xu-Friedman MA & Regehr WG (2004). Structural contributions to short-term synaptic plasticity. *Physiol Rev* **84**, 69–85.

Yi E, Roux I & Glowatzki E (2010). Dendritic HCN channels shape excitatory postsynaptic potentials at the inner hair cell afferent synapse in the mammalian cochlea. *J Neurophysiol* **103**, 2532–2543.

Zagaeski M, Cody AR, Russell IJ & Mountain DC (1994). Transfer characteristic of the inner hair cell synapse: Steady-state analysis. *J Acoust Soc Am* **95**, 3430–3434.

Zampini V, Johnson SL, Franz C, Knipper M, Holley MC, Magistretti J, Masetto S & Marcotti W (2013). Burst activity and ultrafast activation kinetics of $Ca_V1.3$ Ca^{2+} channels support presynaptic activity in adult gerbil hair cell ribbon synapses. *J Physiol* **591**, 3811–3820.

Zampini V, Johnson SL, Franz C, Knipper M, Holley MC, Magistretti J, Russo G, Marcotti W & Masetto S (2014). Fine Tuning of $Ca_V1.3$ Ca^{2+} channel properties in adult inner hair cells positioned in the most sensitive region of the gerbil cochlea. *PLoS One* **9**, e113750.

Zampini V, Johnson SL, Franz C, Lawrence ND, Münkner S, Engel J, Knipper M, Magistretti J, Masetto S & Marcotti W (2010). Elementary properties of $Ca_V1.3$ Ca^{2+} channels expressed in mouse cochlear inner hair cells. *J Physiol* **588**, 187–199.

Zheng J, Shen W, He DZ, Long KB, Madison LD & Dallos P (2000). Prestin is the motor protein of cochlear outer hair cells. *Nature* **405**, 149–155.

Zidanic M & Brownell WE (1990). Fine structure of the intracochlear potential field. I. The silent current. *Biophys J* **57**, 1253–1268.

Additional information

Competing interests

The authors declare no competing interests.

Author contributions

All authors: conception or design of the work; acquisition or analysis or interpretation of data for the work; drafting the work or revising it critically for important intellectual content; final approval of the version to be published; agreement to be accountable for all aspects of the work. All authors have read and approved the final version of this manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All persons designated as authors qualify for authorship, and all those who qualify for authorship are listed.

Funding

This work was supported by an NIH grant RO1 DC014712 to M.A.R. from the National Institute on Deafness and Other Communication Disorders; NIH grants RO1 DC004274 and DC012938 to H.v.G. from the National Institute on Deafness and Other Communication Disorders; and Agencia Nacional de Promocion Cientifica y Tecnologica grant PICT 2016 2155 to J.D.G.

Acknowledgements

K. Clayton, C. Grabner and P. Joris are thanked for valuable comments on an earlier version of the manuscript.

Keywords

first-spike latency, phase-locking, receptor potential, ribbon synapse, synaptic delay, temporal code