Paparella: Volume I: Basic Sciences and Related Principles

Section 2: Physiology

Part 1: Ear

Chapter 8: Electrophysiology of the Peripheral Auditory System

Richard J. Salvi, Flint A. Boettcher, Burnt N. Evans

The purpose of this chapter is to give the reader an overall perspective on how the inner ear transforms the incoming mechanical vibrations into an organized pattern of neural activity that forms the basis for complex auditory perceptions. Of particular interest is how the auditory periphery extracts information about the frequency, intensity, and temporal characteristics of the acoustic stimulus. In order to understand and appreciate many of the important functional characteristics of the inner ear, it may be useful to review some of the basic anatomic features of the cochlea.

Anatomy of the Cochlea

The cochlea consists of a bony tube that spirals about the modiolar axis in the shape of a snail. If the human cochlea were to be uncoiled, it would have a total length from base to apex of approximately 35 mm. Scala tympani, scala vestibuli, and scala media form three fluid-filled chambers within the interior of this bony tube. To appreciate the relationship between these three chambers, it is helpful to unravel the cochlea and view it both longitudinally and in cross-section. Sound information enters the inner ear through the piston-like motion of the stapes, which is loosely coupled in the oval window and in direct contact with the fluids in the scala vestibuli. The fluids in the scala vestibuli and the scala tympani communicate directly with one another through the helicotrema, a small opening near the apex of the cochlea. The basal end of the scala tympani is terminated by a thin membrane known as the round window. The scala media is separated from the other two chambers by Reissner's membrane, the basilar membrane, and bony spiral lamina. The perilymphatic fluid within the scala tympani and the scala vestibuli is similar in composition to extracellular fluid, which has a relatively high concentration of sodium. In contrast, the endolymphatic fluid within the scala media has a relatively high concentration of potassium and is similar to intracellular fluid.

Organ of Corti

The organ of Corti, which lies within the scala media, rests on the osseous spiral lamina and basilar membrane (Fig. 2). The triangularly shaped tunnel of Corti is bounded by the rods of the inner and outer pillar cells. The pillar cells rest on the basilar membrane and lean toward one another so that the pillar heads join at the reticular lamina. On the medial side of the tunnel lie the inner hair cells (IHCs), which are aligned in a single row running from base to apex. Just lateral to the outer pillar cells are three rows of outer hair cells (OHCs). Overlying the organ of Corti is the tectorial membrane, which is hinged medially at the limbus and more loosely attached laterally at the border net above the Hensen cells. Because the basilar membrane and tectorial membrane are hinged at different points, they

move slightly radially with respect to one another when the basilar membrane is deflected upward or downward. The exact anatomic relationship between the tectorial membrane and stereocilia influences the way in which the radial shearing force is transmitted to the hair cell cilia, as discussed below. It appears that the tallest cilia on the OHCs are embedded in the tectorial membrane, since a W-shaped imprint can be seen on the underside of the tectorial membrane (Angelborg and Engstrom, 1973; and Lim, 1980). The IHC cilia, on the other hand, do not appear to contact the tectorial membrane. The OHCs are supported from beneath by Deiters' cells. The thin phalangeal processes of the Deiters' cells project upward to the surface of the organ of Corti, helping to form the tight cell junctions of the reticular lamina. Inner border cells and phalangeal cells surround the IHCs. The Hensen and Claudius cells and the inner sulcus cells provide additional structural support for the organ of Corti.

The basilar membrane increases in width and decreases in thickness from base to apex, while the cross-sectional area of the organ of Corti increases in size (Lim, 1980). These structural variations presumably influence the pattern of mechanical vibration along the cochlear partition. Along the lateral wall of the cochlea lies the stria vascularis, a highly vascularized epithelium that appears to play an important role in regulating the ionic composition and volume of endolymph.

Inner and Outer Hair Cells

Important morphologic differences exist between IHCs and OHCs, suggesting that they may differ functionally. The IHCs, which are approximately 35 microns in length and less than microns in diameter, are pear shaped and contain a centrally located nucleus. The OHCs, in contrast, are cylindrically shaped with a nucleus near the base. OHCs are roughly 6 to 7 microns in diameter, but vary in length from about 25 microns near the base of the cochlea to about 45 microns near the apex (Smith, 1968). The stereocilia on IHCs are arranged in three straight rows, while those on OHCs are arranged in three to four rows in a W-shaped pattern. Within a single hair cell, cilia height increases from the row closest to the modiolus to the row nearest the lateral wall. Ciliary height also systematically increases from base to apex of the cochlea (Lim, 1980). In mammals, each cilium is approximately 0.1 micron in diameter, but tapers considerably as the root enters the cuticular plate. Fine fibrils interconnect the cilia; consequently, when the bundle is deflected, all the hairs are believed to move in unison, pivoting about the root (Flock, 1977; and Pickles et al, 1984). Within the membrane surrounding each cilium is a central core of actin filaments that run in parallel from tip to rootlet. In cross-section, the filaments are hexagonally packed and crosslinked to adjacent filaments at precise intervals, which probably accounts for their great rigidity (Frishkopf et al, 1980; and Saunders and Tilney, 1982). Recent electrophysiologic and microscopic observations from lizards indicate that the length of the cilia may play a role in determining the cell's frequency response characteristics (Holton and Weiss, 1983; and Holton and Hudspeth, 1983).

Afferent and Efferent Innervation

Located within the bony modiolus are the cell bodies of the bipolar spiral ganglion neurons. The axon of the cell projects medially to the cochlear nucleus in the brain stem, while the distal process (or dendrite) projects to the cochlea through the osseous spiral lamina in the organ of Corti. f the 50,000 or so neurons that innervate the cochlea of the cat,

approximately 90 to 95 per cent (type I neurons) synapse directly on IHCs (Fig. 3). Each IHC is contacted at its base by approximately 15 to 20 type I neurons. The remaining 5 to 10 per cent (type II neurons) synapse on OHCs. Each type II neuron branches to innervate approximately 10 OHCs (Spoendlin, 1972).

In the cat, approximately 1800 efferent fibers project to the cochlea from the brain stem (Fig. 4). According to Warr (1978), roughly 63 per cent of these originate from cell bodies located near the ipsilateral superior olivary complex, while the remainder project from cell bodies in the contralateral superior olivary complex. The majority of efferents terminate axodendritically on the afferent dendrites that innervate IHCs. Most of these nerve fibers originate from small fusiform cells located near the lateral superior olive, the majority of which are located on the ipsilateral side of the brain stem. The remaining efferents synapse predominantly on OHCs (axosomatic endings). Most of the efferent fibers innervating OHCs originate from large stellate neurons located near the medial superior olive, the majority of which are located in the contralateral brain stem.

Cochlear Mechanics

Inward motion of the stapes results in displacement of fluid within the scala vestibuli. If this movement were to occur extremely slowly, the fluids in the scala vestibuli would simply be shunted through the helicotrema into the scala tympani, resulting in outward movement of the round window membrane. However, when stapes movement is rapid (> 10 to 20 Hz), it is opposed by the inertia of the fluid mass as well as by the frictional resistance generated by the fluid flow within the narrow scala. This opposition to flow results in a pressure gradient across the basilar membrane. When the stapes is displaced inward, the basilar membrane bulges into the scala tympani, leading to the outward movement of the round window membrane. Rapid inward and outward motion of the stapes thus leads to an alternating pressure gradient that is propagated along the entire length of the cochlea almost instantaneously (Békésy, 1960).

Travelling Wave

The precise characteristics of basilar membrane motion have been a matter of debate dating back before the time of Helmholtz in the 1800s. However, it was not until 1952 that Békésy published the first direct observations of basilar membrane movement using temporal bones from human cadavers (Békésy, 1952). By placing small silver particles on Reissner's membrane and observing their motion with a microscope and stroboscopic illumination, it was possible to estimate the displacement of the basilar membrane using low-frequency tones of high intensity. When the cochlea is stimulated with a pure tone and observed at a particular point, the frequency of basilar membrane motion is the same frequency as the driving signal. If the cochlea is viewed over its entire length, the wave motion would appear to begin at the base and propagate toward the apex, giving the appearance of a "traveling" wave. The peak-to-trough displacement envelope gradually increases with distance from the base, reaching a maximum value at a particular point beyond which the amplitude drops off rapidly. The wavelength (peak-to-peak) of the traveling wave decreases as it moves toward the apex. This indicates that the speed at which the wave propagates toward the apex is decreasing. Consequently, basilar membrane motion near the apex lags behind that in the base. Note that the entire membrane vibrates nearly in phase at very low frequencies and that the phase lag

increases dramatically as frequency increases. The maximum of the velocity envelope is shifted toward the apex relative to that for displacement. The velocity envelope is also narrower near its peak than the displacement envelope.

The top panel shows the *relative* amplitude of basilar membrane displacement for a constant stapes displacement. The amplitude measurements were normalized using the maximum amplitude obtained at each signal frequency; therefore, the peaks of all the curves are the same height. Notice that when the driving frequency is increased, the peak of the traveling wave envelope shifts from the apex toward the base. Thus, the vibration pattern resulting from high-frequency stimuli is primarily confined to the base of the cochlea. Low-frequency tones, on the other hand, can vibrate the entire basilar membrane, but the maximum vibration amplitude always occurs near the apex. In short, the basilar membrane performs a crude spectral analysis on the incoming sound waves, translating frequency onto place.

Frequency Selectivity

To understand how a particular point on the basilar membrane vibrates, one can vary the driving frequency and measure the amplitude of displacement. The most effective frequency for displacing a point 20 mm from the base of the human cochlea is 800 Hz. Only a limited range of frequencies above 800 Hz causes this point to vibrate, whereas any frequency below 800 Hz can cause this point to vibrate if the intensity is high enough. When the data are plotted in this manner, each point on the basilar membrane appears to behave like a bandpass filter; however, the tuning reported by Békésy was much broader than one would expect from psychophysical and physiologic studies.

What physical characteristics of the cochlea determine its vibration pattern? Although this is a complex issue, one can gain an intuitive appreciation of the underlying processes by considering the basilar membrane as a plate that increases in width from base to apex. The increase in width is correlated with a decrease in basilar membrane stiffness by a factor of approximately 100 (Békésy, 1960). When the stapes moves, the pressure is transmitted instantaneously along the entire membrane due to the incompressibility of the fluids. The resulting pressure gradient across the scalae is opposed by the stiffness of the membrane as well as by the inertia of the fluid mass. In a system dominated by mass rather than stiffness, displacement lags behind the driving force. Because of the decrease in stiffness, the apex tends to be mass dominated; therefore, movement in the apex lags behind that in the base. Furthermore, since the resonant frequency of a system is directly proportional to stiffness and inversely proportional to mass, the stiff base responds maximally to high frequencies whereas the compliant apex responds maximally to low frequencies. Movement of the membrane is also opposed by the viscosity of fluids and the frictional forces within the tissue. These forces tend to damp out the motion of the membrane, bringing it to rest within a relatively brief period of time.

Recent Measurements. Békésy's original observations posed a number of serious problems for scientists specializing in hearing because the basilar membrane appeared to behave as a broadly tuned bandpass filter. Physiologic and behavioral measurements, on the other hand, have suggested that the peripheral auditory system is composed of a series of extremely sharply tuned bandpass filters (Evans, 1975; Kiang et al, 1965; Sharf, 1970; and Small, 1959). This discrepancy remained a source of great debate (Evans and Wilson, 1975;

and Geisler et al, 1974) and prompted further experimentation. It should be recalled that due to technical limitations, Békésy's measurements were performed on cadavers at extremely high intensities. More recent experiments carried out on healthy cochleas and using more sensitive measurement techniques (Sellick et al, 1982; and Khanna and Leonard, 1982) have indicated that the basilar membrane is sharply tuned in contrast to Békésy's original experiments.

One set of measurements that demonstrate this sharp tuning were obtained by Sellick and colleagues (1982) using the Mossbauer technique. The measurements were obtained by placing an extremely small radioactive source on the underside of the basilar membrane. The gamma rays emitted by the source were absorbed by a detector with an extremely narrow bandwidth. When the source moves due to vibration of the basilar membrane, the frequency of the radiation changes due to the Doppler shift phenomenon. Thus, the velocity of the basilar membrane modulates the amount of radiation detected by the narrowly tuned absorber, allowing one to estimate its movement. Figure shows the sound intensity needed to produce either a consonant basilar membrane displacement (3.5 Å) or a constant velocity (0.04 mm/sec) for a point near the base of the cochlea. The most effective frequency for exciting this point on the cochlea is 18 kHz, ie, the characteristic frequency (CF). As the driving frequency increases or decreases from the CF, the intensity must be increased significantly in order to produce a criterion velocity or displacement. Thus, the basilar membrane vibration pattern is very sharply tuned in healthy preparations in contrast to the earlier results of Békésy.

Figure shows the growth of basilar membrane velocity as a function of SPL for a point in the cochlea with a CF of 18 kHz (Sellick et al, 1982). The input/output functions are linear below 10 kHz. However, near CF, the input/output functions are linear only at low intensities and exhibit a compressive nonlinearity at higher levels. Consequently, up to a 12-dB change in SPL may be needed to double the vibration amplitude at high intensities. The input/output functions above CF were nonlinear even at low intensities, and the slopes of the input/output functions were shallower than those obtained at low frequencies. One implication of the compressive nonlinearity is that it leads to broader mechanical tuning at high intensities (Sellick et al, 1982). Another interesting aspect of these results is that the compressive nonlinearity gradually disappeared as the physiologic condition of the ear deteriorated (Rhode, 1971; and Sellick et al, 1982).

Temporal Response

The amplitude of most natural sounds varies rapidly over time; therefore, it is of some interest to determine how the basilar membrane responds to transient stimuli. By stimulating the ear with a brief acoustic click, it is possible to examine its damping characteristics. Figure shows the response of the malleus and basilar membrane to a brief acoustic click (Robles et al, 1976). The measurements were obtained in the basal turn of the cochlea using the Mossbauer technique. It is important to note that the middle ear vibrations die out in less than 1 msec whereas the basilar membrane continues to oscillate for more than 3 msec. This indicates that the middle ear is more highly damped than the basilar membrane. In a linear filter, the duration of ringing is inversely related to the bandwidth of the system. Thus, the long duration oscillations suggest that the basilar membrane behaves as a narrowly tuned system. The oscillations in the basilar membrane response also occur at intervals of approximately 0.14 msec, which corresponds to a natural frequency of vibration of roughly

7100 Hz; this is consistent with the frequency predicted from the location of the radioactive source along the cochlear partition.

It is clear from the information presented earlier that the mechanical response of the basilar membrane plays a critical role in analyzing the incoming sound waves. How this pattern of mechanical vibration is transduced into an effective neural code for the central auditory system is the subject of the next section.

Gross Cochlear Potentials

The gross cochlear potentials, which represent the electrical responses summed across many cells, can be grouped into two main categories: those that are elicited with acoustic stimuli versus those that occur in the absence of stimulation. Because the gross potentials are relatively easy to record, they have been used in both basic science and clinical investigations. Unfortunately, they are somewhat difficult to interpret because of uncertainties regarding their generators. Some of the more important characteristics of the gross cochlear potentials are considered in the following discussion.

Endolymphatic Potentials

When an electrode suitable for recording DC potentials is inserted into the scala media and compared against a reference electrode in the scala tympani, one can record a large positive DC potential, known as the endolymphatic potential (EP), in the absence of any acoustic stimulus. Recent studies have found the EP to be approximately + 80 to + 100 mV (Peake et al, 1969; and Bosher and Warren, 1971). The existence of the EP is potentially important, since it could provide a continuous source of energy that could be modulated by a relatively weak acoustic stimulus.

The EP was originally thought to be a diffusion potential resulting from the unique ionic environment of endolymph (high K+ and low Na+ concentration relative to perilymph). However, decreasing the K+ concentration gradient between endolymph and perilymph by injecting KCl into perilymph either had no effect or increased the amplitude of the EP (Konishi and Kelsey, 1968; and Tasaki et al, 1954). Similarly, decreasing the Na+ concentration gradient between endolymph and perilymph by increasing the Na+ concentration in the scala media failed to decrease the EP; instead, it resulted in a slight increase in the EP (Johnstone and Sellick, 1972). It is also important to note that the EP is present only in the endolymph of the cochlear duct and is not seen in the utricle even though the fluids in both compartments are of similar composition (Smith et al, 1958). These results indicate that the EP is not simply a diffusion potential resulting from the ionic composition of the endolymph.

Tasaki and Spyropoulos (1959) helped to establish the source of the EP by removing the fluids from the scala media and probing the cells lining this space. A high positive DC voltage was seen only when microelectrodes contacted the stria vascularis. These findings are consistent with other results showing that the EP is reduced when the stria is damaged surgically (Davis et al, 1958) or with ototoxic drugs such as ethacrynic acid (Sewell, 1984). Thus, the stria is considered by most to be the source of the EP.

The EP appears to require some active metabolic process, since it rapidly declines during anoxia (Johnstone and Sellick, 1972). Immediately after the onset of anoxia the EP shifts to a negative potential, but over the next several hours the potential gradually declines to zero. The negative potential present at the onset of anoxia presumably arises from the diffusion of K+ down its concentration gradient, which leaves behind a net negative charge (Johnstone and Sellick, 1972).

Kuijpers and Bonting (1970) measured the Na+/K+ - dependent ATPase activity in the cochlea and found that it was considerably higher in the stria than in other cochlear structures. Introducing ouabain, which inhibits Na+/K+ ATPase activity, reduced the EP, diminished the endolymphatic K+ concentration, and increased the concentration of Na+ (Konishi and Mendelsohn, 1970; and Konishi et al, 1978). This suggests that the cells of the stria, such as the marginal cells, actively transport K+ into endolymph and extrude Na+ from endolymph against their respective concentration gradients. In order to account for the positive EP, the pump must be electrogenic, ie, more K+ is pumped into the scala media than Na+ is pumped out, resulting in a net positive charge in the scala media. In summary, the EP appears to result from the interaction of two mechanisms: a positive potential generated by the electrogenic pump in the stria vascularis and a negative diffusion potential resulting from the flow of K+ down its concentration gradient.

Cochlear Microphonic

In the 1930s, Wever and Bray performed a rather provocative experiment by placing an electrode on the auditory nerve and listening to the electrical potentials that were generated through a loudspeaker. When one person spoke into the animal's ear, the potentials that were played through the loudspeaker could easily be recognized by another person. This potential was termed the cochlear microphonic (CM).

Origins of the Cochlear Microphonic. The origins and significance of the CM have been the subject of much debate. Wever and Bray (1930) initially felt that the CM originated from the auditory nerve. However, further study showed that the round window was a more effective recording site, and subsequent studies implicated the hair cells as the source of the CM. Tasaki and associates (1954) provided strong supporting evidence for this point of view by measuring the polarity of the CM as a recording electrode passed from below the organ of Corti into the scala media. Passage of the electrode through the basilar membrane into the organ of Corti resulted in a DC voltage drop from 0 to -40 mV, suggesting that the electrode was intracellular. When the electrode entered the scala media, there was an abrupt DC shift from -40 to +75 mV DC. This abrupt DV shift was accompanied by a 180-degree phase reversal of the CM. The polarity reversal strongly suggests that the source of the CM is in the hair cells or some other structure close to the reticular lamina. More recent microelectrode studies have shown that large, sharply tuned microphonic potentials can be recorded from electrodes close to the hair cells (Goodman et al, 1982); this work implicate the sensory cells as the source of the CM. The CM is a graded electrical response that is thought to represent the vector sum (ie both magnitude and phase are important) of the potentials generated by a large number of individual hair cells (Whitfield and Ross, 1965).

Frequency and CM Spatial Distribution. Tasaki and co-workers (1952) used the differential electrode technique to study the distribution of the CM along the cochlear partition

as stimulus frequency was varied. Low-frequency tones resulted in a microphonic response in all three turns of the cochlea. However, as the driving frequency, the microphonic response first decreased in the apical turn and then in the middle turn. Consequently, the CM could only be recorded from the basal turn at the highest stimulus frequency. These results are in general accord with basilar membrane measurements, since they indicate that low-frequency tones are capable of exciting the entire length of the basilar membrane whereas high-frequency tones only produce excitation in the base of the cochlea.

The spatial distribution of the CM can also be examined by plotting the CM voltage at different recording locations, as illustrated in Figure 11. The individual panels show the distribution of the CM at four different frequencies. The curves within each panel were obtained at several intensities; those obtained at the *lowest* SPL (bottom curve in each panel) are the most useful for our purposes, since they reflect the *locally* generated CM. The figure shows that the location of the maximum CM varies with frequency; the maximum is located near the apex at low frequencies and shifts toward the base as frequency increases. In addition, the CM spatial distribution approximates the envelope of the basilar membrane vibration pattern.

Amplitude. One interesting aspect of the CM is that its amplitude remains nearly constant over the duration of the stimulus. This lack of change indicates that the CM does not adapt and is consistent with recent intracellular measurements of the hair receptor potential (Cody and Russell, 1985). The amplitude of the CM, however, does vary with sound intensity. When the logarithm of CM amplitude is plotted against dB SPL, one obtains a relationship similar to that shown in Figure 12 (Dallos, 1973). At low intensities, the amplitude of the CM increases linearly with a slope of approximately one, ie, CM amplitude is proportional to stimulus intensity. As intensity increases further, the amplitude of the CM saturates and then shows a slight decline. At high stimulus intensities, the CM contains a considerable amount of harmonic distortion and may bear little resemblance to the input signal (Dallos, 1973; and Tasaki et al, 1952).

The CM shows no evidence of a threshold; rather the lowest intensity at which the CM can be recorded is limited by the inherent noise in the recording equipment plus biologic noise (Dallos, 1973). Nevertheless, a threshold is often operationally defined, eg, the intensity needed to produce 1 microV of CM. The lowest intensity at which the CM can be reliably recorded depends on the recording location and the stimulus frequency. Electrodes located in the apex are most sensitive to low-frequency sounds, whereas those in the base are most sensitive to high-frequency sounds.

IHC, OHC, and CM Amplitude. One question of considerable practical and theoretical interest is the extent to which the IHCs and OHCs contribute to the CM. Based simply on the fact that there are roughly three times as many OHCs as IHCs, one would expect the OHCs to play a dominant role in generating the CM. Support for this point of view comes primarily from studies in which either the IHCs or OHCs have been selectively destroyed by ototoxic agents (Dallos, 1973; Dallos and Cheatham, 1976). With a complete loss of OHCs, but apparently normal IHCs, the CM input/output function shows roughly as 30- to 40-dB loss in sensitivity and a significant drop in maximum CM voltage. In contrast, CM sensitivity and amplitude are only slightly reduced when the IHCs have been lesioned, but the OHCs are apparently intact. Thus, the OHCs seem to be the dominant source of the

CM.

Mode of Stimulation. Anatomic studies have suggested that the tallest row of stereocilia on OHCs are attached to the overlying tectorial membrane, whereas the cilia on the IHCs appear to be free-standing. If this view is correct, then when the basilar membrane is *displaced,* the relative motion between the tectorial membrane and basilar membrane results in a radial shearing force on the tallest stereocilia of the OHCs. Thus, the OHCs should be sensitive to the *displacement* of the basilar membrane, whereas the free-standing cilia on the IHC should be stimulated by the viscous fluid drag of the endolymph that occurs during basilar membrane motion (*velocity*). To test this hypothesis, Dallos and Durrant (1972) examined the CM waveform in response to triangular motion of the stapes. Stapes velocity is *constant* during the triangular wave, except at its peak and trough where there is an abrupt change in velocity (ie acceleration) and a reversal in the direction of motion. A constant stapes velocity results in a constant pressure at the oval window, a constant *displacement* of the basilar membrane, and a constant CM amplitude. Since the CM is generated primarily by OHCs in normal ears, one must conclude that the OHCs are responding to basilar membrane displacement. What, then, do the IHCs respond to? To examine this issue, OHCs in the first turn were eliminated with ototoxic drugs, leaving behind apparently intact IHCs. In this case, triangular displacement of the stapes resulted in a CM response at the peak and trough of the triangular wave, ie, during stapes acceleration. Thus, stapes *acceleration* leads to a brief pressure change at the oval window and in basilar membrane motion (velocity) and a transient increase in CM amplitude. In summary, the cochlea has two groups of receptors: OHCs capable of responding to basilar membrane displacement and IHCs sensitive to its velocity (Dallos et al, 1972; Dallos, 1973).

Model of CM Production. A number of different models have been used to explain the CM, but one which accounts for much of the data is Davis' variable resistance-battery model, which schematized in Figure 13. The EP provides a +80-mV battery for driving current through the reticular lamina into the scala tympani. The driving force behind the current flow is further amplified by the negative intracellular potential of the hair cells, which acts like a battery in series with the EP. The flow of current is normally limited by the high-resistance of the reticular lamina. When the stereocilia are bent away from the modiolus, the resistance of the hair cells is presumably decreased, resulting in an increased current flow. Conversely, when the ciliae are bent toward the modiolus, resistance is increased, leading to a decrease in current flow. Thus, the change in hair cell resistance brought about by basilar membrane motility modulates the current flow across the organ of Corti and results in a voltage fluctuation corresponding to the CM.

Summating Potentials

When an acoustic stimulus is initiated, one can record a DC potential shift from intracochlear electrodes in addition to the CM. This DC shift, known as the summating potential (SP), appears to have multiple origins. The waveforms shown in Figure 14 (Dallos et al, 1970) serve to illustrate the complex nature of the SP. Note that the DC potentials recorded in the scala tympani or the scala vestibuli can be either positive or negative and that they can reverse polarity depending on the stimulus frequency, level, and duration. Dallos and co-workers (1970) have derived two measures of the SP that seem to behave in an orderly manner when stimulus frequency and recording location are varied. To obtain these measures,

the potentials recorded from the scala tympani and the scala vestibuli were averaged to eliminate the CM and isolate the DC components (Fig. 14). As illustrated in Figure 15, the DC voltages were subtracted to obtain the difference component (DIF = ST - SV), which reflects the output of generators located between the recording electrodes in the scala tympani and the scala vestibuli. The potentials were also added together to obtain the average component (AVE = (ST + SV)/2). The AVE response was previously thought to reflect only remote electrical activity that appeared in phase at both electrodes. However, under certain conditions, it may also represent a locally generated response having the same polarity in both the scala tympani and the scala vestibuli (Dallos et al, 1970; Dallos, 1973).

The left panel of Figure 15 (Dallos, 1975) shows the amplitude of the DIF component plotted as a function of frequency for electrodes placed in the first, second, and third turns of the cochlea. For an electrode located in the third turn, the DIF component is positive (DIF+) at low frequencies, but as frequency increases it reverses to a negative potential (DIF-) at some best frequency. Note that the best frequency of the DIF systematically increases as the recording location shifts from apex to base. The AVE plots in the right panel of Figure 15 behave in a complementary manner. At each recording location, the AVE is positive (AVE+) over a narrow band of frequencies and negative (AVE-) above or below this frequency band. Note that the frequency band producing the AVE+ is similar to the band that produces the DIF-. It is interesting that the best frequency for the AVE+ and DIF are tuned approximately two-thirds of an octave higher in frequency than the peak of the CM and the basilar membrane displacement envelope, ie, they are located on the high-frequency slope of the traveling wave.

The DIF+. DIF-, and AVE+ components of the SP, as well as the CM, do not seem too vary in amplitude as stimulus off-time is reduced. The lack of adaptation in these three components of the SP suggests that they may share a common generator with the CM, namely the hair cells. The AVE-, on the other hand, decreases in amplitude and may even reverse polarity when stimulus off-time is shortened; the adaptation-like properties of the AVE are, therefore, similar to those of the auditory nerve. This suggests that the AVE may reflect postsynaptic activity, perhaps the generator potential from the nonmyelinated region of the auditory nerve. Results obtained from animals with missing OHCs have provided further insights into the source of the SP. The DIF-, DIF+, and AVE+ appear to be generated primarily by OHCs at low to moderate intensities. At higher SPLs, the IHCs contribute significantly to DIF- and AVE+ but not to DIF+ (Dallos, 1975).

Compound Action Potential

When the hair cells are activated, they presumably release a neural transmitter that depolarizes the afferent dendrites and causes an all-or-none spike discharge in individual auditory nerve fibers. A gross electrode placed in the vicinity of the cochlea may be used to sample the activity from these neurons provided that they are activated with a high degree of synchrony. The method of achieving the required synchrony is to give an acoustic stimulus that has a relatively abrupt onset such as a click. The resulting neural response, known as the compound action potential (AP), is graded in amplitude and represents the sum of the unit responses added in various phases and amplitudes (Antoli-Candela and Kiang, 1978). As illustrated in Figure 16, the AP waveform consists of two prominent negative peaks, the N1 and N2, which occur approximately 1.0 and 2.0 msec after the onset of the stimulus. N1

almost certainly represents the response of units in the auditory nerve. The origins of N2 are less certain; some believe that units in the cochlear nucleus contribute significantly, whereas others believe that it represents secondary discharges of units in the auditory nerve. Approximately 0.7 msec prior to the N1 response, one can also observe a small deflection in baseline activity that represents the CM; the delay between CM and AP is characteristic of that imposed by a synapse. The AP appears to be of neural origin since it rapidly disappears during anoxia, whereas the CM persists although reduced in amplitude.

Origin of the AP. The high degree of neural synchrony necessary to elicit the AP is most easily achieved near the base of the cochlea, where mechanical travel time is short, leading to near simultaneous activation. As the vibration pattern moves toward the apex, the travel time progressively increases, resulting in a loss of neural synchrony between adjacent segments of the cochlea. Consequently, the AP response is weighted toward the activity of neurons located in the base of the cochlea. In spite of this problem, it is possible to obtain information from neurons in the apex of the cochlea if low-frequency tone pips are presented at low intensities.

Amplitude and Latency. Figure 17 (Salvi et al, 1979, 1983) illustrates how the amplitude and latency of N1 vary with stimulus intensity. In this case, the response was elicited with a 0.1-msec click; 0 dB attenuation of this signal corresponds to a peak SPL of approximately 100 dB. Note that as stimulus intensity increases, N1 grows in amplitude over a 40- to 50-dB range before saturating and rolling over. The input/output function for N1 parallels that for the CM; however, the AP response can usually be recorded at slightly lower stimulus intensities than the CM (Davis et al, 1958). In hearing-impaired persons, the input/output function tends to be shifted to the right, reflecting a loss in sensitivity as illustrated in Figure 17. Furthermore, there may be a reduction in maximum amplitude.

As shown in Figure 17B, when the stimulus level is increased, the time between the onset of the stimulus and the peak of N1 decreases. At intensities where the N1 is just detectable (80-dB attenuation or 20-dB SPL), the latency is approximately 2 msec. As intensity is increased, the latency decreases rapidly over the first 40 to 50 dB, followed by a more gradual decline to a plateau at roughly 1 msec. The shape of the AP latency-intensity function can be altered substantially by cochlear impairment; however, the specific change depends on the nature of the hearing loss. The open symbols shown in Figure 17B illustrate the type of latency-intensity function seen in animals with a relatively flat sensorineural hearing loss. Note that the latency is quite short at the intensity where it can first be detected. The latency-intensity function is also quite shallow compared to normals. Such information may prove useful in differential diagnosis.

AP Versus Behavioral Threshold. Recent studies have shown that the N1 response is relatively frequency specific at low intensities and can serve as a reasonable predictor of the behavioral audiogram (Dallos et al, 1978). To illustrate this point, Figure 18 compares the average behavioral threshold of the chinchilla with the mean AP threshold and the mean single unit thresholds of auditory nerve fibers. In this case, the AP "threshold" was defined as the intensity needed to produce a noticeable N1 response (1 microV). The important point to note here is that the AP threshold curve parallels the behavioral audiogram over much of the frequency range; however, it is shifted upward by approximately 10 to 20 dB. A similar correlation between the AP and the behavioral threshold has been observed in human studies

of electrocochleography, indicating that the AP may be useful for estimating the audiogram of difficult-to-test subjects (Eggermont and Odenthal, 1974).

Frequency Selectivity. One of the hallmarks of the auditory system is its frequency selectivity, it, the ability to extract or respond to a particular frequency component in a complex signal. One method of estimating the frequency selectivity of the peripheral auditory system is through the use of AP tuning curves (Dallos and Cheatham, 1976). To obtain an AP tuning curve, a fixed-frequency, low-level tone burst (probe tone) is presented 10 to 15 dB above threshold in order to elicit a small but consistent AP response from a restricted region of the basilar membrane. Then a second continuous tone, known as a masker, is introduced and increased in intensity until it just abolishes the AP response to the probe tone (ie the masked threshold). When the masked threshold is determined over a range of masker frequencies above and below the probe, one obtains a V-shaped tuning curve. As shown in Figure 19, the masked threshold is lowest when the masker frequency is equal to the probe frequency and systematically increases as the frequency separation between masker and probe increases. Masking presumably occurs because the masker invades the region of the cochlea stimulated by the probe tone, resulting in a loss of neural synchrony to the probe tone. The general features of the AP tuning curve are remarkably similar to those of the basilar membrane mechanical response (see Fig. 7). This similarity in shape suggests that the frequency selectivity of the AP response may be derived, to a large extent, from the inherent frequency selectivity of the basilar membrane.

Hair Cell Physiology

Over the past decade, researchers have made significant advances in understanding the physiologic properties of individual hair cells (Crawford and Fettiplace, 1980; Russell and Sellick, 1978; Tanaka et al, 1980; and Weiss, 1974). By directing extremely fine microelectrodes into the organ of Corti, it is possible to record intracellularly from individual hair cells and to mark the recording location by dye injection for later histologic analysis.

Inner Hair Cells

In 1977, Russel and Sellick successfully recorded from IHCs in the basal turn of the guinea pig cochlea. The resting potentials of these cells ranged from -35 to -45 mV, and both an AC and a DC response were recorded in response to sound as illustrated in Figure 20 (Sellick and Russell, 1978). The upper trace of each pair represents the amplitude of the DC response of the hair cell to a tone burst, while the lower trace represents the amplitude of the AC response. At low stimulus levels, small AC and DC responses occur to a narrow range of frequencies near the characteristic frequency (CF) of the cell. However, as stimulus intensity increases, the amplitude of both the AC and the DC responses increases and the response patterns broaden, particularly toward the low frequencies. The DC response reaches a maximum amplitude of approximately 12 mV at the most sensitive frequency, whereas the AC response is usually less than 0.6 mV. The AC response also decreases with frequency so that the ratio of AC to DC amplitude declines at the rate of 6 to 9 dB/octave. This frequency-dependent attenuation presumably occurs because the AC component is shunted across the membrane capacitance of the hair cell.

Amplitude-Intensity Functions. Figure 21 shows the amplitude-intensity functions for the AC and DC receptor potentials at different stimulus frequencies. Both the AC and DC amplitude-intensity functions increase linearly at low intensities, but begin to saturate at higher SPLs. The amplitude-intensity function at CF is located furthest to the left and saturates at lower SPLs since the cell is most sensitive to this frequency. Since the cell is less sensitive to frequencies above or below CF, the intensity functions are shifted further toward the right as the frequency separation between the driving frequency and CF increases. The AC and DC amplitude-intensity functions obtained above CF tend to saturate sooner and at lower amplitudes than those obtained at lower frequencies; this frequency effect parallels that seen in basilar membrane vibration input/output functions obtained above CF (Sellick et al, 1982).

Membrane Resistance. The Davis variable resistance-battery model depicted in Figure 13 assumes that a decrease in the resistance at the top of the hair cell increases the current flowing from the scala media into the hair cell, resulting in depolarization. Sellick and Russell (1978) have measured the hair cell membrane resistance as well as the membrane depolarization while presenting acoustic stimuli. The resistance of the hair cell decreased as stimulus intensity increased; furthermore, the decrease in membrane resistance was linearly related to the amount of hair cell depolarization, as illustrated in Figure 22. The slope of the curve was approximately 0.62 mV/megaohm. Thus, the magnitude of the hair cell receptor potential was directly related to the membrane resistance of the hair cell as predicted by the model.

Frequency Selectivity. One important issue is the extent to which the frequency selectivity of the IHC is correlated with the frequency response of the basilar membrane. One method of estimating hair cell frequency selectivity is to determine the intensity needed to produce some criterion DC voltage that is just detectable (threshold) above the baseline noise of the recording system. When such measurements are made over a range of frequencies, one obtains a tuning curve similar to that shown in Figure 23 (Sellick et al, 1983). The tip of the tuning curve, or CF, is located at 18 kHz. "Threshold" increases extremely rapidly above and below CF; however, the IHC continues to respond to a wide range of low-frequency tones, but only at high intensities. For comparison purposes, one can also determine the intensity necessary to produce a constant basilar membrane displacement or velocity at a point in the cochlea corresponding to the location of the IHC. It can be seen that the isodisplacement and isovelocity contours follow the IHC isopotential curves fairly closely, suggesting that the frequency selectivity seen in the IHC is derived to a large extent from basilar membrane motion. However, it is unclear from this data whether displacement, velocity, or some combination of the two is the critical stimulus variable.

The evidence presented so far suggests that the frequency selectivity of the hair cell is derived principally from the mechanical input to the stereocilia rather than to some inherent property of the hair cell itself. However, in more primitive cochleas (eg lizards), mechanical and neural tuning are often poorly correlated; therefore, other mechanisms for tuning have been sought. In some species, tuning appears to be closely related to cilia length such that the cells with the longest cilia have receptor potentials with the lowest CFs. Direct observations have shown that low-frequency stimuli produce the largest cilia displacements on hair cells having the longest cilia, whereas high-frequency stimuli only cause displacement on hair cells with short cilia (Holton and Hudspeth, 1983). In this case, the tuning appears to be derived principally from the mechanical properties of the cilia.

In addition to the mechanical tuning derived from basilar membrane displacement and cilia length, there is also evidence that some tuning occurs *following* the mechano-electrical transduction process (Crawford and Fettiplace, 1980; 1981; and Hudspeth, 1985). When a hair cell from a turtle cochlea or frog sacculus is depolarized by a constant current step, one finds that the membrane potential undergoes a damped, sinusoidal oscillation before reaching a stable potential. The frequency of the damped oscillation corresponds closely to the cell's resonant frequency to mechanical stimulation. Figure 24 illustrates how membrane resonance is thought to occur. The initial opening of the transduction channels by mechanical stimulation allows for the influx of K+ into the hair cell. The resulting depolarization causes voltage-sensitive Ca2+ to open, leading to the influx of Ca2+, which further depolarizes the cell. As the Ca2+ concentration increases, Ca2+-sensitive K+ begins to open, leading to the efflux of K+ and repolarization of the cell. Thus, the damped membrane potential appears to result from the interplay of inward Ca2+ current and outward K+ current (Lewis and Hudspeth, 1983).

Hair Cell Polarization. Based on indirect results from CM recordings, it was suggested that IHCs respond to basilar membrane velocity, whereas OHCs respond to basilar membrane displacement. To provide direct evidence for this contention, Sellick and Russell (1980) recorded from IHCs during trapezoidal displacement of the basilar membrane, as shown in Figure 25. *Displacement* of the basilar membrane toward the scala vestibuli deflects OHC cilia away from the modiolus. This results in a constant, positive CM during each half-cycle of the stimulus from an electrode beneath the reticular lamina. Since OHCs are the dominant source of the CM, a positive CM would be indicative of depolarization of the OHCs. As the basilar membrane moves (*velocity*) toward the scala vestibuli, a large depolarization occurs in the IHC; this depolarization is absent during static displacement of the basilar membrane. Conversely, during basilar membrane motion toward the scala tympani, there is a slight hyperpolarization of the IHCs. These results suggest that the OHCs are activated by basilar membrane displacement toward the scala vestibuli and that the IHCs are depolarized by basilar membrane velocity toward the scala vestibuli and hyperpolarized by velocity toward the scala tympani.

Efferent Stimulation Effect. The majority of efferents in the crossed olivocochlear bundle (COCB) terminate on OHCs; however, a small proportion also end on the efferents to the IHCs. Several studies have shown that activation of the COCB during sound stimulation leads to an increase in the amplitude of the CM but a reduction in the output of the APP and single auditory nerve fibers (Fex, 1962; and Wiederhold, 1970). The reduction in auditory nerve output is surprising when one considers that only a few COCB efferents terminate on afferent neurons that innervate IHCs. In order to better understand the underlying mechanisms, Brown and colleagues (1983) stimulated the COCB while recording from IHCs. As illustrated in Figure 26, COCB stimulation had no effect on the IHC amplitude-density function obtained with tone bursts below the cells' CF. However, it reduced the amplitude of the receptor potential to tones presented at CF; this effect was prominent at low intensities and gradually disappeared at high intensities. Because of the frequency-dependent nature of the reduction, COCB activation resulted in an elevation of the tip and a slight broadening of the IHC tuning curve. It is important to note that COCB stimulation did not alter the resting potential of the cell or alter its membrane resistance; thus there is no evidence that COCB stimulation *directly* affects the IHC. Instead, the COCB may *indirectly* influence the IHCs through their action on OHCSs; changes in OHCs could potentially alter basilar membrane

mechanics and thereby indirectly influence the response of IHCs (Mountain, 1980; Siegel and Kim, 1982).

Outer Hair Cells

OHCs have proved to be much more difficult to study than IHCs. The initial evidence (Tanaka et al, 1980) indicated that OHC resting potentials were rather small. However, more recent work (Dallos et al, 1982) indicates that the OHCs have large resting potentials (-71 mV) and AC receptor potentials up to 15 mV peak to peak.

Frequency Response. Panel A in Figure 27 shows the amplitude of the AC potential at 40 dB SPL as a function of frequency in an OHC and IHC near the apex of the cochlea. All the functions show bandpass characteristics and exhibit a peak near 800 Hz consistent with the expected best frequency for this recording location. The frequency response areas for IHCs and OHCs are fairly similar except that the low-frequency tail of the tuning curve is somewhat shallower. The extracellular AC potential, by comparison, is less sharply tuned. Panel B in Figure 27 provides a comparison of the AC and DC tuning characteristics of an OHC at a higher intensity, 70 dB SPL. The first point to note is that the DC potential is more sharply tuned than the AC response. Second, as intensity increases, the peak of the AC response shifts to a lower frequency but the DC does not (compare panel A obtained at 40 dB to panel B obtained at 70 dB).

IHCs and OHCs share a number of common features. First, the frequency response of OHCs is approximately the same as that of IHCs near CF. Second, both IHCs and OHCs exhibit roughly the same sensitivity. Third, at *low intensities* there is no systematic detuning, or difference in CF for IHCs and OHCs obtained from the same recording location (Honrubia et al, 1976; and Zwislocki, 1975). Although there are some functional similarities between IHCs and OHCs, there are also some important differences. First, the OHC resting potential is approximately twice as large as that from an IHC; however, the OHC AC and DC receptor potentials are approximately one-third the size of those recorded from IHCs (Cody and Russell, 1985; and Dallos et al, 1982). Second, hyperpolarizing responses are more prevalent in OHCs than IHCs. For example, Dallos and associates (1982) reported that OHCs in the apex display hyperpolarizing DC responses to stimuli below CF, whereas IHCs always depolarize. This reversal in polarity with frequency is similar to that seen for the extracellular SP (see Fig. 19). Furthermore, the AC responses from OHCs in the basal turn become asymmetric in the hyperpolarizing direction at high intensities, whereas the AC response from IHCs become asymmetric in the depolarizing direction (Cody and Russell, 1985). Finally, there is some debate regarding the phase relationship between the AC response from IHCs and OHCs.

Transduction. The electrical signals that occur when the cilia are deflected arise from the flow of ions through channels in the cell's membrane. The location of the transduction channels has been inferred by measuring the voltage gradients in the fluids surrounding isolated hair cells from the sacculus of the bullfrog (Hudspeth, 1982). The largest voltage gradients occurred near the tips of the hair bundle, suggesting that the transduction channels are localized to that region. By substituting different ions in the bathing medium, it has been possible to characterize the transduction channels. K+, which is normally present in the extracellular fluid, readily passes through the membrane along with other small alkali and

small organic cations. These results indicate that the channel has a diameter of approximately 0.7 nm (Corey and Hudspeth, 1979; and Hudspeth, 1985).

When the cilia are deflected, electrical responses can be detected in hair cells within a few microseconds. Since the response is extremely rapid, the transduction channels may be opened directly by mechanical tension in the cilia rather than involving some intermediate transduction step. A model of transduction which incorporates many of the preceding findings is illustrated in Figure 28 (Hudspeth, 1985). The transduction channels, which are believed to be located in the tips of the cilia, are coupled to the next tallest row of cilia by a flexible link. An anatomic structure similar to the proposed link has been recently identified in specially prepared electron micrographs (Pickles et al, 1984). When the cilia bundle is deflected toward the tallest cilia, tension is created in the transduction link, forcing a greater proportion of channels to remain open and increasing the influx of cations. When the bundle is deflected in the opposite direction, tension in the link forces a greater proportion of the channels closed.

OHC Motility. New techniques for tissue dissociation using enzymes or mechanical trituration have allowed researchers to study sensory receptors and supporting cells from the inner ear in vitro (Fig. 29). Initial work in this area has revealed that OHCs are motile and not just simply passive transducers (Brownell et al, 1985; and Evans et al, 1986). Two different types of movement have been observed in OHCs. The first type (type I) can be elicited by electrical stimulation and consists of a cycle-by-cycle modulation in cell length around some mean resting length. When OHCs are artificially depolarized by intracellular current injection, the cells become shorter and wider whereas hyperpolarizing currents cause OHCs to become longer and narrower. The change in length is distributed between the nucleus and the cuticular plate region but becomes increasingly restricted to the apical region of the OHC as the stimulating frequency increases. Transcellular electrical stimulation also results in contractile activity. Stimulation is most effective when one electrode is placed near the cuticular plate and the other is near the synaptic pole (Fig. 29). Initial attempts at measuring the frequency response of type I motion has revealed relatively low-pass characteristics with maximum displacements of less than a micron (Brownell et al, 1985; and Evans et al, 1986). However, in more recent studies, frequency following responses has been observed up to 8 kHz (Ashmore and Brownell, 1986). With improvements in instrumentation, it may be possible to measure OHC displacements down to the nanometer range and out to higher frequencies. Brownell (1986) has proposed that type I movements may be mediated by an electro-osmotic mechanism.

A second type of motile response (type II) with a long time constant has also been observed. Brownell and associates (1985) and Evans and colleagues (1986) has shown that a slow *increase* in cell length is superimposed upon the faster type I movement when low-intensity electrical pulse trains are used to stimulate OHCs. However, at high levels of stimulation, there is a slow *decrease* in overall length. Type II responses also occur with the introduction of various pharmacologic agents. When acetylcholine or cholinomimetic analogues are applied to the synaptic pole. OHCs decrease in length. Application of these compounds to other sites along the cell failed to produce any motion (Brownell, 1986; and Evans et al, 1986).

What cellular structures and processes could be involved in OHC motility? Zenner (1986) identified actin not only in the cuticular plate and stereocilia but also below the cuticular plate down to the base of the cell. Myosin, which is required in any actin-mediated contraction, was also identified near the cuticular plate and along the cell membrane extending toward the base of the cell. Cytochalasin, which inhibits the polymerization of actin filaments and actin-dependent contractions, suppressed the contractile response of OHCs. Using detergent-treated OHCs with increased membrane permeability, Zenner (1986) reported contractions when Ca2+ and ATP were introduced into the external bathing medium. The degree and rate of contraction depended on the adenosine triphosphate concentration. The contractile response was suppressed if trifluoperazine, a substance which inhibits the calcium-binding protein calmodulin, was added to the bathing medium. The response was also absent if Ca2+ was absent from the bathing medium; however, if inositol triphosphate, a component of membrane phospholipids, was added to the medium the response was reestablished. This suggests that inositol triphosphate might act as a second messenger, causing the release of Ca2+ sequestered within the cell, perhaps in the subsurface cisternae (analogous to the sarcoplasmic reticulum of skeletal muscle) lining the plasma membrane. Evans and co-workers (1986) examined the subsurface cisternae in OHCs during successive stages of OHC contraction (Fig. 30). As the cell begins to contract, the most medial cisternae begin to bud off into adjoining vesicles. If a strong contraction is elicited, the cisternal system progressively rearranges into a multivesiculated matrix.

Brownell (1986) has suggested that the laminated cisternal system may be involved with the fast type I contractile response seen with electrical stimulation. Electro-osmosis is thought to provide the driving mechanism for this model of OHC motility, ie an electric field forces the cloud of ions adjacent to a charged membrane to migrate, which, in turn, forces the fluid adjacent to the charge membrane to move. The fluid flow presumably occurs between the plasma membrane and the cisternal system and results in a pressure gradient, and therefore movement, along the long axis of the OHC. The appeal of the model is that the response would be extremely fast and proportional to the applied voltage. Furthermore, the CM could provide the current for driving an electro-osmotic mechanical response. Finally, the relaxation that follows a strong contraction may be brought about by the high turgidity of the OHCs, which would force the cell to return to its normal length.

What possible role can movement of OHCs have on cochlear function? Some insights into the functional significance of the mechanical response can be gleaned from the anatomic schematic (Brownell et al, 1985) shown in Figure 31. The base of each OHC sits in a cup-like structure formed by the Deiters' cells. The phalangeal process of the Deiters' cell projects toward the cuticular plate of an adjacent hair cell, aiding in the formation of the reticular lamina. According to the scheme proposed by Brownell and co-workers (1985), the triangular structure formed by the cuticular plate, the OHC, and the phalangeal process of the Deiters' cell would become more rigid as the OHC lengthened and, conversely, more compliant if it shortened. This, in turn, would influence the compliance of the cochlear partition, thereby altering its vibration characteristics to acoustic stimuli. Such a scheme would be compatible with the effects of COCB stimulation on the IHC receptor potential (Brown et al, 1983). Furthermore, a change in length of the OHCs could alter the position of the overlying tectorial membrane, since the tallest stereocilia on the OHCs appear to be embedded in it. Finally, the electromechanical response may provide feedback to cochlear partition motion. Although the preceding results strongly suggest that there is some active mechanical process in the cochlea,

it is not yet clear whether or not the active process involved in OHC motility is responsible for the great sensitivity and sharp mechanical tuning of the cochlear partition.

Cochlear Emissions

Traditionally, the transduction process has been thought to be unidirectional, i.e., mechanical to neural. However, the recent discovery of acoustic emissions emanating from the ear coupled with evidence for mechanical responses from OHCs has argued strongly for transmission in the reverse direction, physiologic to mechanical.

Acoustically Evoked Emission

In 1978, Kemp showed that when the human ear is stimulated with a click, two distinct acoustic events could be recorded in the ear canal, as illustrated in Figure 32. The first acoustic component, recorded within the first 6 msec, is due to the impulse response of the sound source, ear canal, middle ear, and passive components of the cochlea. As expected, the first component increases linearly with sound intensity. However, 8 to 9 msec after the click, a secondary acoustic emission (Kemp echo) occurs, which is smaller in amplitude (generally < 25 dB SPL). The amplitude of the secondary emission grows linearly at low intensities, but a compressive nonlinearity is initiated at high levels; consequently, roughly a 3-dB increase in sound intensity is needed to produce a 1-dB increment in the level of the emission. The secondary emission is clearly biologic in origin since it is reduced or abolished by ototoxic agents (Kemp, 1978, 1982), whereas the initial acoustic event is unaffected.

The spectrum of the click-evoked emission varies over time. Near onset, the emission contains mostly high-frequency energy, but the spectrum gradually shifts toward the low frequencies later in time. Tone pips can also be used to evoke the emission. In this case, however, the spectrum of the emission is the same as the tone. The latency of the emission increases as stimulus frequency decreases in much the same way as the basilar membrane travel times increase with decreasing frequency. However, the emission latency is somewhat longer than the estimated time it would take for a signal to propagate to its characteristic place in the cochlea and back into the ear canal. This suggests that some intervening step is involved in emission generation.

If two tones of similar frequency are used to stimulate the ear, the examiner can detect distortion products in the acoustic emissions. The intermodulation distortion products are most pronounced at the frequencies 2F1 - F2, 2F2 - F1, and F2 - F1. Siegel and Kime (1982) showed that the amplitude of the distortion products could be altered by electrically stimulating the fibers of the COCB, most of which synapse on OHCs. Since infusion of curare into perilymph blocked the effect of COCB stimulation, it is reasonable to conclude that the process is synaptically mediated. These observations, coupled with the observed motility of OHCs, suggest that the central nervous system is capable of altering cochlear biomechanics via the OHC system.

Spontaneous Emissions

All of the evoked emissions discussed in the previous sections required some external acoustic signal to be detected. However, as early as 1960, there were a few clinical reports

of spontaneous ear canal emissions intense enough to be heard by individuals standing near the patient (Glanville et al, 1971; and Huizing and Spoor, 1973). Unfortunately, these results went largely unnoticed until the discovery of the Kemp echo. With the aid of a sensitive microphone and a spectrum analyzer, it has been possible to identify spontaneous acoustic emissions (Fig. 33) in approximately one-third of all normal human ears (Zurek, 1981, 1985). Since few subjects can hear their own emissions, it seems unlikely that spontaneous emissions are a significant cause of tinnitus.

Spontaneous emissions are typically less than 20 dB SPL and generally occur in narrow bands located below 2 kHz. Spontaneous emissions remain stable within an ear but vary in amplitude and spectrum across ears. Using external tones, it is possible to suppress spontaneous emissions. Tuning curves for spontaneous emission can be determined by finding the frequency-intensity combinations that produce a criterion reduction in the amplitude of the emission. These spontaneous emission tuning curves are similar in shape to those seen for basilar membrane vibration (see Fig. 11). The presence of these spontaneous acoustic emissions strongly supports the view that there is a biologically active process that supplies mechanical energy to the cochlea.

Auditory Nerve Physiology

The auditory nerve, which consists of 30,000 to 50,000 neurons, is the only pathway for relaying information from the cochlea to the central auditory system (Harrison and Howe, 1974). As shown in Figure 3, approximately 90 to 95 per cent of the neurons innervate IHCs, each IHC, however, is contacted by approximately x afferents, making it a highly redundant system (Spoendlin, 1972, 1978). Since the majority of nerve fibers innervate single IHCs, the response patterns observed from a single fiber reflect the output of an extremely small region along the basilar membrane. By recording from many nerve fibers in a single animal, it is possible to gain an overview of the neural input to the central auditory system. Recordings from single units in the auditory nerve have typically been carried out by inserting microelectrodes into the root of the auditory nerve as it exits the internal auditory meatus; however, some investigators have also been able to record from cells in the spiral ganglion (Robertson and Johnstone, 1979).

Spontaneous Activity. In the absence of any controlled acoustic stimulation, most auditory nerve fibers discharge spontaneously. When recordings are obtained from a large sample of units, one finds that the majority of nerve fibers have spontaneous rates ranging between 0 and 100 spikes per second, as shown in the histogram in Figure 34 (Salvi et al, 1983). The distribution of spontaneous rates tends to be bimodal, with one large, narrow peak near 0 spikes/sec and a second, broader peak situated between 40 and 70 spikes/sec (Kiang et al, 1985; and Liberman, 1978). Few units seem to have spontaneous rates around 20 spikes/sec. Since the temporal patterns of spontaneous activity is highly irregular, it is common to evaluate such data by obtaining an interspike interval histogram which shows the distribution of the time intervals between successive spike discharges. The interspike interval histograms contain few counts at extremely short intervals owing to neural refractoriness. This is followed by a peak located at 10 msec or less. The decay from the peak is approximately exponential. Thus, when the ordinate is scaled logarithmically, as in Figure 35, the decay from the mode approximates a straight line (Kiang et al, 1965).

Threshold and Frequency Selectivity. If tone bursts of the proper frequency and intensity are presented to the ear, one can detect an increase in the unit's discharge rate above the spontaneous rate. By adjusting the intensity at each frequency so that a just detectable (threshold) increase in firing rate occurs, one can map out the unit's tuning curve (Kiang et al, 1965). Figure 36 shows a series of tuning curves obtained from six different auditory nerve fibers. Each of the V-shaped tuning curves has a narrowly tuned, low-threshold region. The frequency at the tip of the tuning curve is referred to as the unit's CF. Threshold increases significantly above or below CF. Units with high CFs also exhibit a high-threshold, broadly-tuned, low-frequency tail. The shapes of the auditory nerve fiber tuning curves are remarkably similar to the isoamplitude response curves of the basilar membrane and the IHC receptor potential.

Threshold and Spontaneous Activity. When recordings are obtained from a large sample of units in a single animal, one finds a systematic relationship between spontaneous activity and threshold among units with similar CFs, as illustrated in Figure 37 (Liberman, 1978; and Salvi et al, 1982). Units with high spontaneous rates (> 18 spikes/sec) have the lowest thresholds and comprise roughly 66 per cent of the population. Units with medium spontaneous rates (1 to 18 spikes/sec) comprise 23 per cent of the population and have thresholds that are approximately 7 dB higher than the low-threshold units. The remaining 11 per cent of the units have very low spontaneous rates (< 1 spike/sec) and high thresholds. Some low-spontaneous rate units have thresholds that may be up to 50 dB higher than the most sensitive units; however, the tips of their tuning curves are as narrow as those with the lowest thresholds.

IHC and OHC Afferents. By recording intracellularly from auditory nerve fibers and then injecting horseradish peroxidase into a cell, it is possible to relate the physiologic properties of primary afferents with their pattern of innervation on IHCs and OHCs. The results of these studies indicate that low and medium spontaneous rate fibers have thin dendrites that terminate on the side of the IHC facing the modiolus (Liberman, 1980; 1982; and Robertson, 1984). The high spontaneous rate fibers, on the other hand, have thick dendrites, which terminate on the side of the IHC facing the OHCs. The synapses from the thin and thick fibers are similar in many respects; however, some of the thinnest fibers have a larger and more complex synaptic membrane specialization and a longer synaptic bar. It is important to note that all of the labeled fibers that *responded to sound* innervated IHCs; none contacted OHCs (Liberman, 1982).

Robertson (1984) recorded intracellularly from another group of cells that had steady negative resting potential but failed to generate action potentials either in quiet or in the presence of intense broad-band noise. Only 1 of the 17 "silent" neurons was successfully labeled. The peripheral processes of this cell projected radially across the tunnel of Corti, then spiraled basalward for approximately 400 microns before giving off a series of small endings, each terminating at the base of an OHC. The lack of spike activity in this OHC afferent suggests that little or no information may be conveyed to the central auditory system along this pathway; instead, most spike activity is transmitted by way of the IHC afferents.

Firing Patterns. When a tone burst is presented above threshold, a unit responds at a high rate over the duration of the tone; however, the discharge pattern varies in a probabilistic manner from one stimulus presentation to the next. To characterize the average

firing pattern, it is necessary to repeatedly present the stimulus and sample the discharge pattern over many trials to obtain a post-stimulus time histogram. Figure 38 shows a series of post-stimulus time histograms collected over a range of sound intensities using a 200-msec tone burst at the unit's CF (3470 Hz). Near threshold, the number of spike discharges is just barely detectable above the spontaneous rate and the post-stimulus time histogram is nearly flat. As intensity increases, the overall number of spikes increases; the firing rate is highest near stimulus onset and declines to a quasi steady-state level over the remainder of the stimulus. The decay in firing rate is characterized by a relative fast component within the first 4 to 6 msec, followed by a more gradual decay over the next 40 to 50 msec (Smith, 1980). It is important to note that the CM and hair cell receptor potentials do not show a similar decay in output over time. Therefore, the neural adaptation seen in the response of single auditory nerve fibers most likely results from the properties of the hair cell-nerve fiber synapse (Furukawa and Matsuura, 1978).

Recovery. The discharge rate of a neuron not only depends on the acoustic stimulus but also on the unit's previous history of stimulation, ie, the system has short-term memory (Harris and Dallos, 1979; and Smith, 1977, 1978). One method of demonstrating the short-term memory of the system is to measure a unit's firing rate to a short-duration, low-level probe tone. When the probe tone is presented in isolation, the unit responds in a robust manner. However, when the probe tone is preceded by an adaptor tone of long duration and high intensity, the response to the probe tone is greatly reduced. As illustrated in Figure 39, the reduction in the probe tone discharge rate is greatest at short adaptor-probe intervals. As the time interval between the adaptor tone and probe tone increases, the unit's firing rate increases so that by approximately 200 to 300 msec it is completely recovered. This recovery in discharge rate is also thought to involve the hair cell-nerve fiber synapse.

Phase Locking. When stimulated by low-frequency tones (< 5 kHz), auditory nerve fibers tend to respond to preferred points within the stimulus cycle, a phenomenon referred to as phase locking. The neural discharges do not necessarily occur on every cycle but may skip a beat due to neural refractoriness. One particularly useful method of evaluating phase locking involves sorting the spike counts with respect to their time of arrival within the stimulus period (Rose et al, 1971). When the spike counts are collected over many stimulus cycles, one obtains a period histogram similar to those shown in Figure 40. The solid line in Figure 40 represents the sinusoid signal that was fit to the histogram. It is important to note that the spikes tend to occur only during one-half of the stimulus cycle, ie, the response is a half-wave rectified version of the stimulus. This presumably occurs because deflection of the stereocilia away from the modiolus leads to depolarization of the hair cell (see Fig. 25) and the release of neurotransmitter that activates the nerve fiber. Deflection of the cilia toward the modiolus leads to hyperpolarization and a lack of transmitter release. In mammals, phase locking is lost at frequencies above 4 to 5 kHz. One factor that may account for this is temporal jitter in the spike initiation mechanism. Another possibility suggested by Russell and Sellick (1978) is the attenuation of the AC receptor potential at high frequencies. When the AC component is lost, the DC potential leads to a steady release of neurotransmitter and a loss of phase locking.

Input/Output Functions. Figure 41 illustrates how the firing rate of a unit varies with stimulus intensity. At the CF of 16.6 kHz, there is a systematic increase in firing rate from approximately 40 to 80 dB SPL, after which the firing rate appears to plateau (Sachs and

Abbas, 1974). The input/output functions obtained at frequencies below CF show a similar increase in firing rate with level, but the rate-intensity functions are shifted to the right since the unit is less sensitive to frequencies below CF. The input/output functions above CF are also shifted to the right, but the slopes of the rate-intensity function are reduced. This decrease in slope above CF is similar to that seen for the response of the basilar membrane and the IHC receptor potential (see Figs. 8 and 21).

The rate-intensity functions at CF generally increase rapidly within the first 20 to 30 dB of threshold. In the majority of fibers, further increases in intensity cause little or no increase in firing rate so that the input/output function displays a flat saturation rate. However, a significant proportion of fibers exhibits a sloping saturation; ie, the firing rate increases rapidly over the first 20 to 30 dB, followed by a more gradual increase over a range of 30 dB or more (Sachs and Abbas, 1974). Liberman (1978) has reported a relationship between the saturation discharge rate and spontaneous activity. Units with low spontaneous rates have lower saturation discharge rates than units with medium and high spontaneous rates. The saturation discharge rate also increases with CF, particularly units with CFs greater than 4 kHz (Liberman, 1978).

When very high intensities are used, many units with CFs less than 2 kHz show a dramatic decline in discharge rate around 90 dB SPL, followed by an abrupt increase to the plateau discharge rate when intensity is increased further (Fig. 42). In some cases, the dip in the discharge rate-intensity can approach the spontaneous discharge rate. Furthermore, as intensity approaches 90 dB, the preferred phase of firing abruptly shifts by 180 degrees. The abrupt change in both rate and phase have suggested to some that there may be two components contributing to the rate-intensity functions of auditory nerve fibers, one operating at low intensities and a second dominating at high levels (Gifford and Guinan, 1983; Kiang et al, 1986; and Liberman and Kiang, 1984).

Response to Clicks. An acoustic click has an extremely broad energy spectrum and fast rise time; therefore, it is especially useful in synchronizing the activity of single auditory nerve fibers that contribute to the AP. To understand how the AP is generated, it is useful to examine the discharge patterns of auditory nerve fibers to click stimuli (Antoli-Candela and Kiang, 1978). The post-stimulus time histograms shown in Figure 43 illustrate the type of response patterns that are obtained from single fibers with different CFs. The histograms from high CF units contain a single peak that occurs approximately 2 msec after click onset. The post-stimulus time histograms obtained from low CF fibers contain multiple peaks, and the latency to the first peak may be 3 msec or longer. A illustrated in Figure 44, the latency to the first peak is correlated with CF. Units with CFs greater than a 4-kHz-range have latencies between 1 and 1.5 msec. Below 4 kHz, the latency increases with decreasing CF. The time interval between the histogram peaks (DP) can also be used to estimate the CF of low-frequency units by the relationship CF = 1/DP.

Based on the fact that the hair cells are polarized, one would predict that the temporal structure of the post-stimulus time histogram would be dependent on click polarity, ie, whether the stapes was initially pulled out (rarefaction) or pushed in (condensation). When histograms are collected to both condensation and rarefaction clicks, as in Figure 45, one finds that the peaks to rarefaction clicks occur slightly earlier in time than those for condensation clicks (Kiang et al, 1965; and Pfeiffer and Kim, 1972). The shorter latency to

rarefaction stimuli is consistent with the view that hair cell excitation is initiated by basilar membrane motion toward the scala vestibuli. In order to appreciate the temporal structure of the neural response, one can combine the histogram to rarefaction clicks with an inverted version of the histogram to condensation clicks to form a compound histogram. The decaying peaks in the compound post-stimulus time histogram are remarkably similar in shape to the damped sinusoidal vibration or ringing seen in the basilar membrane vibration curves (see Fig. 9). Thus, the neural response is similar to what one would expect from a narrowly tuned filter. Energy from the broad-band click is extracted around the center frequency of the filter, and the sharp tuning results in a damped sinusoidal response corresponding to the center frequency of the filter.

Nonlinear Responses. If the auditory system processed acoustic signals in the same way as a linear system, then the response to two tones would simply be the sum of the responses to the individual components. The output of the auditory nerve, however, exhibits a number of nonlinearities. One significant nonlinearity that has already been mentioned is the half-wave rectification or locking of neural discharges to a particular half of the stimulus cycle. Other nonlinearities exist, the most prominent of which are two-tone suppression and combination tone responses.

Two-tone suppression can be readily demonstrated by stimulating a unit with a continuous tone at CF (F1) and then presenting a second tone (F2) just above or below F1 (Kiang et al, 1965; and Salvi et al, 1982). When F1 is presented alone at an intensity above threshold, the firing rate increases above the spontaneous rate, as shown in Figure 46. A similar increase in firing rate can be seen when F2 is present alone. Based on these results, one might expect that presentation of F1 and F2 together would result in a discharge rate greater than to either stimulus alone. However, as can be seen in Figure 46, the simultaneous presentation of F1 and F2 results in a firing rate that is less than that for either tone alone and that is temporarily less than the spontaneous discharge rate.

Arthur and associates (1971) mapped out the frequency-intensity combination of F2 that would produce a reduction in firing rate for an F1 tone located at CF. The frequency regions leading to suppression occurred along the high- and low-frequency borders of the tuning curve, as illustrated in Figure 47. More recent studies have shown that the suppression areas can occur at other frequency regions if F1 is displaced from CF. Thus, the location of the suppression are, it seems to depend on the relative position of F1 and F2 and their relationship to the tuning curve.

One aspect of the data that may provide clues to the underlying physiologic mechanism is the latency of two-tone suppression. Since suppression effects occurs nearly instantaneously, this would tend to rule out the efferent system, which would involve a relatively long time delay. Furthermore, sectioning the efferent input to the cochlea failed to eliminate suppression (Kiang et al, 1965). Sellick and Russell (1979) demonstrated that suppression can occur as peripherally as the IHC. The filled circles in Figure 48 show the excitatory response area (tuning curve) for the DC potential of an IHC obtained with a single tone. A 17-kHz test tone (F1) was then used to stimulate the IHC at its CF and a second tone (F2) was introduced and varied in both frequency and level. The open triangles in Figure 48 show the intensity of F2 needed to produce a 20 per cent *reduction* in the DC receptor potential to the 17-kHz tone. The suppression areas for the IHC DC response flank the

excitatory response area and are remarkably similar to the auditory nerve fiber suppression areas. These results suggest that the nonlinear response could be occurring in the hair cells or basilar membrane response. Rhode (1977) has reported some evidence of two-tone suppression in the mechanical response of the basilar membrane; however, it is not yet clear if this can completely account for all the neural data.

Another nonlinearity that has been demonstrated in the auditory nerve is a response to combination tones. Under the proper conditions, an auditory nerve fiber responds only when two tones are presented simultaneously; however, if either tone is presented alone, the unit may not respond to either (Goldstein and Kiang, 1968). The response to two-tone stimulation is most robust when F1 and F2 are chosen such that the intermodulation tones (F2-F1) and (2F1-F2) correspond too the unit's CF. Given that this information can be conveyed through the auditory nerve, it should not come as a surprise that human listeners can perceive the distortion tones F2-F1 and 2F1-F2. Although combination tone responses have been detected in basilar membrane motion, their amplitudes appear to be far below those needed to account for the psychophysical and physiologic data (Rhode, 1977; and Wilson and Johnstone, 1973).

Over the past quarter century, enormous progress has been made in understanding the mechanisms by which acoustic energy is transformed into a coherent neural code that can be effectively utilized by the central auditory pathway. Still, there is much to be learned about the important steps in this chain of events. During the next decade, we can expect to see rapid growth in our understanding of the biophysical processes responsible for the sharp mechanical-neural tuning of the basilar membrane, the cellular processes involved in transduction, neurotransmission, and the neural cording mechanism involved in processing complex stimuli such as speech.