═══REVIEW COMMONS MANUSCRIPT═══

IMPORTANT:

- Manuscripts submitted to Review Commons are peer reviewed in a journal-agnostic way.
- Upon transfer of the peer reviewed preprint to a journal, the referee reports will be available in full to the handling editor.
- The identity of the referees will NOT be communicated to the authors unless the reviewers choose to sign their report.
- The identity of the referee will be confidentially disclosed to any affiliate journals to which the manuscript is transferred.

GUIDELINES:

- For reviewers: https://www.reviewcommons.org/reviewers
- For authors: https://www.reviewcommons.org/authors

CONTACT:

The Review Commons office can be contacted directly at: office@reviewcommons.org

Main Manuscript for

Slow Mechanical Filtering by Outer Hair Cells Enhances Rather than Limits Receptor Potential Kinetics.

Henrik Sahlin Pettersen[1,2], Anders Fridberger[3] and Pierre Hakizimana[3]*

1. Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway.
2. Department of Pathology, St. Olav's hospital, Trondheim University Hospital, Trondheim, Norway.
3. Department of Biomedical and Clinical Sciences (BKV), Linköping University 581 83, Linköping, Sweden.

* To whom correspondence and requests for materials should be addressed.
Email: pierre.hakizimana@liu.se

Author Contributions: A.F. developed the automated data acquisition, aligning it with the established acquisition paradigm and contributed with expertise in cochlear mechanics. P.H. conceptualized and designed the study, as well as conducted the experimental procedures. The segmentation of the hearing organ into structures of interest was a joint effort between P.H. and H.S.P., with H.S.P doing more work. The analysis of the neural networks' output data was a joint effort by H.S.P. and P.H., working from datafiles generated by H.S.P. P.H. took charge of analyzing the electrical recordings and performing statistical quantifications. The manuscript was written by P.H., with input from the co-authors.

Competing Interest Statement: The authors declare no competing interests.

Classification: Biological Sciences / Neuroscience.

Keywords: Electromotility, Adaptation, Cochlear amplification, AI-enhanced rapid microscopy.

This PDF file includes:

Main Text
Figures 1 to 5
Tables 0
Supplementary Figures 1 to 4

A Supplementary Information file is also provided with 1 to 136 tables

Abstract

Mammals achieve extraordinarily sensitive hearing through the active amplification of sound vibrations in the cochlea. Central to this process is the electromotility of outer hair

cells (OHCs)—their somatic length changes powered by the membrane motor protein prestin, which converts receptor potentials into mechanical forces.

Yet the intrinsic RC (resistance–capacitance) low-pass filtering of the OHC membrane poses a paradox: how can OHCs generate the rapid force changes necessary for effective amplification? One hypothesis suggests that hair bundle adaptation, operating on sub- to tens-of-millisecond timescales, might compensate for the comparatively slower somatic response.

In this study, we resolve the paradox by simultaneously recording sound-evoked electrical and mechanical responses in the guinea pig cochlear apex during adaptation-inducing intense sound stimulation. High-speed confocal microscopy paired with AI-based segmentation enabled two-dimensional quantification of the OHC mechanical response—tracking changes in length, width, and area—while extracellular recordings captured receptor potentials.

Under control conditions, the inherently slow kinetics of prestin-mediated mechanical changes serve to sharpen receptor potential kinetics. Notably, pharmacological blockade of prestin with salicylate reversed this relationship: the mechanical response kinetics became markedly sharper, while receptor potential kinetics shifted into a sluggish, low-pass regime.

These findings demonstrate that the intrinsic low-pass filtering of electromotility is not a limitation but rather a critical feature that preserves rapid electrical transduction, thereby underpinning the exquisite sensitivity of the mammalian auditory system.

Significance Statement

Understanding the mechanisms underlying cochlear amplification is central to deciphering the exquisite sensitivity of mammalian hearing. Our study reveals a reciprocal interplay between mechanical and electrical filtering in outer hair cells. Specifically, we demonstrate that the intrinsic low-pass filtering of OHC electromotility is essential for sharpening receptor potential kinetics. When prestin is pharmacologically blocked, the mechanical response kinetics become markedly sharper while receptor potential kinetics become low-pass filtered —thereby reversing the normal interplay. By combining high-speed confocal imaging with AI-based segmentation and extracellular recordings in the guinea pig cochlear apex, our findings resolve a longstanding paradox in auditory biophysics. These insights significantly deepen our understanding of cochlear amplification and may inform novel therapeutic strategies for hearing impairments.

Main Text

Introduction

Mammals achieve remarkably sensitive hearing through *active* amplification of incoming sound vibrations in the cochlea—a process essential for speech perception, music appreciation, and survival. Central to this amplification are the outer hair cells (OHCs),

which contribute forces via two principal mechanisms: somatic electromotility[1–6] and hair bundle motility[7–15]. When acoustic waves reach the cochlea, they deflect OHC hair bundles[7,16–21], opening mechanoelectrical transducer (MET) channels that allow ion influx and generate the receptor potential[16,22–33]. This voltage drives OHC somatic electromotility, wherein the motor protein prestin underlies voltage-dependent changes in cell length[2,3,6,34–39]. By enhancing basilar membrane motion, OHCs boost sound sensitivity by orders of magnitude[2,3,6,34–39].

While the term electromotility was originally used to describe electrically evoked OHC length changes[3,6], it has evolved to encompass voltage-dependent contractile responses of OHCs regardless of how the voltage changes are initiated. In the intact cochlea, sound-induced hair bundle deflection leads to receptor potentials that drive these contractile responses through prestin—a mechanism recently demonstrated in vivo by Dewey et al. (2021) [40], who observed sound-evoked antiphasic motions between the top and bottom of OHCs at frequencies exceeding 20 kHz in mice. Following their demonstration, we use this terminology interchangeably with the sound-evoked OHC mechanical changes observed in our measurements, which include changes in cell length, width, and area.

A longstanding puzzle, however, concerns the membrane's RC (resistance–capacitance) properties, which ostensibly low-pass filter voltage at high frequencies[38,40–42]. In principle, this should limit how effectively OHCs can generate rapid force at frequencies critical for mammalian hearing. As a result, some have proposed that hair bundle adaptation (including sub-millisecond force generation at the hair bundles) might dominate high-frequency amplification, relegating somatic electromotility to lower frequencies[38,41,42]. Yet other evidence—including OHC motion measured above 20 kHz[40,43]—suggests that somatic electromotility remains important across the mammalian hearing range, even though its basic electrical properties appear "too slow."

To reconcile this paradox, we focus here on relatively long-timescale (tens-to-hundreds of milliseconds) DC responses in living OHCs[31,44,45] that occur alongside the much faster (sub-millisecond) adaptation of MET channels[38,41,42,46]. These DC responses, though slower than the nominal RC cutoff, are known to *scale* with stimulus intensity[31,44,45] and may critically shape how receptor potentials are ultimately tuned. Despite extensive study of OHC electromotility in isolated cells[1–6], establishing its precise role *in situ* has remained challenging. Methods that directly correlate electrical and mechanical signals within the mostly intact cochlea (albeit without efferent feedback)[45,47–52]—where all supporting cells, fluids, and membranes remain physiologically coupled—are needed.

Here, we developed a new approach combining (i) high-speed confocal microscopy[47] with AI-based segmentation for 2D mechanical quantification[53], (ii) simultaneous extracellular recordings of the OHC receptor potential[45,51], and (iii) targeted pharmacological manipulations of key molecular pathways. Specifically, we blocked prestin with salicylate to isolate its contribution to electromotility[4,54], modulated extracellular calcium with ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EGTA)[45,51] to probe MET-channel adaptation, and partially blocked MET channels with dihydrostreptomycin (DHS)[45,50,55,56]. While our recordings come from the guinea pig apical turn (best frequency of ~180 Hz[45,51])—a region more experimentally accessible for

confocal imaging[47,49,51,52,57]—these mechanisms are fundamental to OHC function throughout the cochlea.

Strikingly, we find that the slow mechanical filtering inherent to OHC electromotility actually *sharpens* the receptor potential kinetics rather than limiting it. Blocking prestin reverses this relationship—mechanical response kinetics become faster, but receptor potential kinetics shifts to a low-pass regime. These results suggest that "low-pass" OHC electromotility is, in fact, critical for preserving higher-frequency content in the receptor potential. Our findings resolve how electromotility remains effective even above its nominal RC cutoff: by filtering slower voltage components and maintaining sharp transduction kinetics. Beyond clarifying this core biophysical puzzle, these insights highlight a novel principle of mechanical–electrical co-filtering in sensory systems, offering a compelling explanation for the remarkable frequency selectivity and sensitivity of mammalian hearing.

Results

We developed an integrated experimental platform that combines precision acoustics, electrical recording, and high-speed confocal imaging to investigate the relationship between direct current (DC) outer hair cell (OHC) electromotility and receptor potential kinetics in a nearly intact cochlear apex[45,47–52] (Suppl. Fig. 1a-b). Using the guinea pig temporal bone preparation[45,47–52], we maintained the cochlea in near in vivo conditions through gentle perfusion with oxygenated cell culture medium (see Methods for detailed procedures; see Ref[57] for an image of the preparation). Sound stimuli were delivered through the intact ear canal and synchronized with extracellular recordings obtained via an electrode positioned near the organ of hearing, accessed through the intact Reissner's membrane (see Ref[45] for a depiction of the electrode placement near the hearing organ's imaging site). Mechanical responses were visualized using reflection-mode confocal microscopy (Zeiss LSM 980, 488 nm), enabling rapid acquisition of 512×512-pixel images within 630 ms without the need for fluorescent labeling[47].

To characterize DC electromotility and receptor potential dynamics, we delivered 160 Hz sound bursts (near the location's best frequency of ~180 Hz[45,51]) at 105 dB SPL during sustained stimulation. Each burst comprised a 720 ms active segment, whose intensity varied depending on the experimental phase (105 dB SPL during sustained stimulation and 0 or 70 dB SPL during control periods, Suppl. Fig. 1a), bracketed by 140 ms silent intervals to enable baseline DC response (also known as summating potential, SP[31,45]) measurements. Confocal imaging windows (630 ms), synchronized with the center of the active segments, captured DC mechanical responses, while continuous 2.5-minute recordings tracked receptor potential dynamics during sustained stimulation (Suppl. Fig. 1b).

The protocol included four one-minute control periods—two preceding (0 and 70 dB SPL) and two following (70 and 0 dB SPL) the sustained stimulation. These control periods were designed to establish mechanical baselines at 0 dB SPL and electrical baselines at 70 dB SPL, capturing distinct aspects of baseline activity under both unstimulated and moderate acoustic conditions (Suppl. Fig. 1b). This approach ensured

robust comparisons of mechanical and electrical responses across baseline (0 dB SPL), moderate stimulation (70 dB SPL), sustained stimulation (105 dB SPL), and recovery (post-stimulation periods at 70 and 0 dB SPL).

Electromotility Block Shifts Receptor Potential Kinetics to a Low-Pass Regime

Figure 1a shows high-temporal resolution electrical recordings at three key timepoints during a 412-s stepped acoustic stimulation protocol (0 → 70 → 105 → 70 → 0 dB SPL): t_1 = 93 s (mid-70 dB SPL), t_2 = 209 s (mid-105 dB SPL), and t_3 = 318 s (recovery 70 dB SPL). At t_1, the response exhibits a large AC (green) component and a positive SP (blue) relative to baseline. By t_2, the AC is attenuated while the SP becomes strongly negative. Notably, fast adaptation is evident in the SP signal—with roughly 60% of its decrease occurring within the 140-160 ms window (≈6.4 cycles at 160 Hz)—and slow adaptation appears as gradual SP amplitude changes over the 159-s sustained stimulation period (Figs. 1a–b). Upon returning to 70 dB SPL at t_3, the electrical response nearly recovers, though the SP remains slightly negative (Figs. 1a–b).

The averaged temporal profile (Fig. 1b) reveals a biphasic adaptation: the SP rapidly drops to ~–0.8 mV within 57 s and then enters a plateau phase, during which its timing remains closely linked to the AC. Recovery at 70 dB SPL is notably faster for the SP than for the AC (Fig. 1c). Importantly, the AC response's dependence on stimulus intensity stays robust across pharmacological treatments—including salicylate, EGTA, and DHS (Fig. 1d).

Fast adaptation dynamics were quantified by first-order exponential fits (using a consistent automated procedure; see Methods, Fig. 1g and Supplementary Information (SI), Table 1) of the initial 6.4 cycles at t_2. Under control conditions, the fit yielded a time constant τ = 0.09 ms, amplitude = 0.62 mV, and offset = –0.62 mV (R^2 = 0.948, n = 24; Figs. 1e–g). A brief oscillation between 146–150 ms was observed (insert in Fig. 1e; also visible in Fig. 1g) but decayed rapidly.

To probe underlying mechanisms, targeted pharmacological interventions were applied. Local delivery of 100 µM EGTA (via a pre-loaded glass microelectrode into the endolymphatic space near the TM and hair bundles[45,47,50,51]) lowered extracellular Ca^{2+} and thereby reduced calcium entry through MET channels. EGTA preserved the adaptation amplitude (0.61 mV, offset = –0.64 mV, R^2 = 0.968, n = 12) but prolonged τ to 0.14 ms—highlighting calcium's role—while the 146–150 ms oscillation persisted (indicating calcium independence).

Similarly, local application of 5 µM DHS (delivered identically[45,50]) reduced K^+ entry via MET channels. DHS produced kinetics similar to EGTA (τ = 0.14 ms) but lowered the adaptation amplitude to 0.36 mV (offset = –0.44 mV, R^2 = 0.953, n = 13), with the 146–150 ms oscillation again persisting, suggesting its independence from potassium modulation.

In contrast, treatment with 10 mM salicylate (added to the gently perfused, oxygenated cell culture medium[57]; see Methods) produced the most pronounced changes: fast

adaptation slowed five-fold (τ = 0.46 ms) and the amplitude was reduced to 0.20 mV (offset = –0.22 mV, R^2 = 0.895, n = 12), with adaptation completing within 140.4–145.7 ms time window (SI, Table 1). Notably, salicylate abolished the 146–150 ms oscillation, linking that feature to electromotility.

Linear mixed model analysis during the fast adaptation window (140–180 ms; SI, Tables 2–3) confirmed that—relative to control—salicylate significantly slowed adaptation (estimate = 0.48 ± 0.11, t = 4.44, $p < 0.001$, F = 19.72) while DHS trended toward slower kinetics (estimate = 0.17 ± 0.10, t = 1.79, p = 0.074, F = 3.19). EGTA did not significantly affect the rate (estimate = 0.009 ± 0.12, t = 0.08, p = 0.939, F = 0.006). Significant time-dependent terms ($p < 0.001$) across conditions further support robust adaptation (F values: salicylate = 1145.86; EGTA = 5828.99; DHS = 4196.27).

Slow adaptation was analyzed over a 57-s SP response window (126–183 s, Figs. 1h–j). In controls, slow adaptation had an amplitude of 1.10 mV, offset of –1.11 mV, and τ = 0.01 s (R^2 = 0.948, n = 24; SI, Table 4j). EGTA reduced the amplitude to 0.64 mV (offset = –0.86 mV) and prolonged τ to 0.19 s (R^2 = 0.953, n = 12), confirming calcium-dependent regulation. DHS maintained near control kinetics (τ = 0.02 s) but reduced the amplitude to 0.68 mV (offset = –0.70 mV, R^2 = 0.927, n = 13), whereas salicylate further attenuated the amplitude to 0.34 mV (offset = –0.37 mV) and slowed kinetics to τ = 0.30 s (R^2 = 0.985, n = 12).

Statistical analyses of stimulus-intensity effects during high-intensity (105 dB SPL) windows (Figs. 1k–t; SI, Tables 5–14) reveal distinct treatment-specific patterns. In controls, the AC amplitude dropped from 2.41 ± 0.15 mV to 1.43 ± 0.08 mV (p = 0.001) during 105 dB SPL, with only partial recovery (1.36 ± 0.13 mV, p = 0.001 vs. baseline; Fig. 1k). Salicylate further enhanced this attenuation (from 1.71 ± 0.15 mV to 1.01 ± 0.09 mV during 105 dB SPL, then to 0.92 ± 0.06 mV at 70 dB SPL; Fig. 1l). EGTA preserved the initial AC response (1.99 ± 0.41 mV vs. 1.51 ± 0.15 mV, p = 0.424) before a delayed drop (1.17 ± 0.17 mV, p = 0.043; Fig. 1m), while DHS maintained initial amplitude (1.19 ± 0.17 mV vs. 1.25 ± 0.14 mV, p = 0.500) before a later reduction (0.93 ± 0.12 mV, p = 0.017; Fig. 1n).

Mixed-effects modeling over the 126–183 s window (evaluating slow adaptation time constants; Fig. 1; SI, Tables 9–11) showed that salicylate had the strongest suppressive effect on AC responses (estimate = –0.45 ± 0.02, $p < 0.001$, t = –22.04), followed by DHS (estimate = –0.21 ± 0.02, $p < 0.001$, t = –10.52), with EGTA showing no significant change (estimate = –0.00 ± 0.02, p = 0.819, t = –0.23). Time effects were minimal (estimate = 0.0004 ± 0.0004, p = 0.302, t = 1.03) and ANOVA confirmed a highly significant treatment effect ($F(3,3533) = 215.21$, $p < 0.001$).

For the SP amplitude, controls showed a robust negative deflection during 105 dB SPL (from 0.29 ± 0.06 mV to –0.85 ± 0.10 mV, p = 0.001) with partial recovery (–0.11 ± 0.04 mV, p = 0.001; Fig. 1p; SI, Tables 7–8). Salicylate nearly abolished this deflection (0.33 ± 0.11 mV to –0.04 ± 0.14 mV, p = 0.034; Fig. 1q), while EGTA (Fig. 1r) and DHS (Fig. 1s) maintained a negative shift (EGTA: 0.11 ± 0.07 mV to –0.71 ± 0.14 mV; DHS:

0.10 ± 0.04 mV to –0.59 ± 0.15 mV, both $p = 0.001$) with partial recovery (–0.17 ± 0.06 mV and –0.06 ± 0.02 mV, respectively). Further mixed-effects modeling (SI, Tables 12–14; Fig. 1t) indicated that salicylate exerted the strongest modulation (estimate = 0.71 ± 0.02, $p < 0.001$, $t = 33.01$), followed by DHS (estimate = 0.12 ± 0.02, $p < 0.001$, $t = 5.78$) and EGTA (estimate = 0.04 ± 0.02, $p = 0.039$, $t = 2.07$), with an overall significant negative time trend (estimate = –0.004 ± 0.000, $p < 0.001$, $t = –10.78$; ANOVA: $F(3,3533) = 465.16$, $p < 0.001$).

In summary, our data reveal critical interactions between electromotility and receptor potential kinetics. Under control conditions, fast adaptation ($\tau \approx 90$ µs) falls within the established range (50–150 µs) [33,58,59], whereas EGTA prolongs τ to 140 µs—suggesting calcium dependence, though the difference is not statistically significant. DHS reduces adaptation amplitude without altering τ, whereas salicylate markedly slows fast adaptation ($\tau \approx 460$ µs, $p < 0.001$), highlighting a strong coupling between prestin-mediated electromotility and receptor potential kinetics. Similarly, slow adaptation is slowed[60] by EGTA ($\tau = 190$ ms) and most profoundly by salicylate ($\tau = 300$ ms, $p < 0.001$). These treatment-dependent profiles support a mechanistic framework in which prestin and calcium signaling modulate rapid, sustained cochlear responses, prompting studies on prestin and cochlear mechanics.

Outer Hair Cell Electromotility Sharpens Receptor Potential Kinetics

To investigate the temporal dynamics of outer hair cell (OHC) receptor potential kinetics and electromotility, we used an advanced image analysis framework that combines confocal microscopy with deep learning–based segmentation[53] (Fig. 2a). In each experiment (acquisition pipeline, Suppl. Fig. 1), 412 confocal images were captured and processed via a two-stage denoising protocol followed by expert annotation[53]. Quality control retained only high-quality time series, discarding few low-contrast datasets (3/24 in controls, 4/12 in EGTA, 3/13 in DHS, and 0/12 in salicylate).

We implemented DeepLabV3 with ResNet18 (Fig. 2a, Methods) using an active learning paradigm that iteratively incorporates expert-verified predictions and challenging cases. Post-processing included adaptive thresholding, morphological refinements, multi-layer quality control, metric computation, and automated outlier detection.

Our segmentation achieved high accuracy with IoU scores of 0.87–0.94 and pixel accuracy >0.98 across seven cochlear structures: OHCs (IoU = 0.90), TM (IoU = 0.94), HeCs (IoU = 0.87), RL (IoU = 0.88), PCs (IoU = 0.90), IHC (IoU = 0.91), and BCs (IoU = 0.92) (Figs. 2b–f; Supplementary Movies 1–2; see Methods).
Post-segmentation, the framework computes length, width, area, and centroid coordinates for each structure to capture OHC electromotility and other structural changes (radial and axial displacements). Baseline correction subtracts the median of the first 63 data points at 0 dB SPL. Thus, the baseline-corrected centroids directly indicate radial/axial displacements (negative values denote movement toward the spiral limbus and basilar membrane respectively), and variations in length, width, and area reflect mechanical responses (negative values signify compression).

The dataset—comprising seven segmentation classes, five metrics, four pharmacological conditions (including controls), 412 data points per time series, and five stimulus intensity phases—is extensive. Therefore, the main text emphasizes key insights, with supplementary material providing additional details. Each dynamic metric is overlaid with the corresponding average electrical recording (Fig. 1i) to ensure consistent visual reference.

OHC length and width changes were measured during sound stimulation (Fig. 2g; Suppl. Movie 3). At 105 dB SPL onset, OHCs exhibited a decaying exponential response with a length change of –3.53 ± 1.9 µm and a width change of –2.81 ± 0.6 µm (n = 21). During the 126–183 s sustained phase, time constants were $\tau = 60$ ms ($R^2 = 0.864$) for length and $\tau = 50$ ms ($R^2 = 0.844$) for width (Suppl. Fig. 2; SI, Table 15). Notably, the electrical responses (Fig. 1h–j) were markedly faster.

Pharmacological interventions further probed these dynamics. With salicylate (Fig. 2h; SI, Table 15), electromotility accelerated: length changes had $\tau = 20$ ms ($R^2 = 0.855$, amplitude 5.49 µm) and width changes $\tau = 40$ ms ($R^2 = 0.763$, amplitude 2.01 µm; note that the width R^2 is below the 0.8 threshold). These are considerably faster than the receptor potential kinetics ($\tau = 300$ ms), indicating that lowpass filtering of the electrical response enhances electromotility.

Under extracellular micromolar EGTA (Fig. 2i; SI, Table 15), electromotility was severely attenuated (length and width dynamics with $R^2 < 0.8$), although receptor potential kinetics remained measurable ($\tau = 190$ ms). With extracellular micromolar DHS (Fig. 2j; SI, Table 15), length changes showed moderate kinetics ($\tau = 50$ ms, $R^2 = 0.898$, amplitude 4.51 µm) while width changes were slower ($\tau = 450$ ms, $R^2 = 0.825$, amplitude 2.08 µm), alongside accelerated electrical responses ($\tau = 20$ ms).

Controls in mechanically distinct structures further validate these findings. PCs (Fig. 2k–n; SI, Table 15), known for high stiffness[61], exhibited no reliable length or width adaptation ($R^2 < 0.8$) under control conditions or with salicylate, EGTA, or DHS. In contrast, the TM (Fig. 2o–r; SI, Table 15), with gel-like compliance[62], expanded by 2–3 µm in both dimensions; however, adaptation remained below reliable thresholds ($R^2 < 0.8$), precluding exponential fitting. TM expansion was particularly pronounced with EGTA (Fig. 2q), consistent with its established calcium-dependent properties[47,51,63] and suggesting that TM-associated calcium is critical for both auditory function and structural integrity.

Area measurements demonstrate that two-dimensional analysis yields superior resolution of OHC electromotility compared to one-dimensional metrics (Figs. 3a–c, Suppl. Fig. 3; SI, Tables 15–16; Suppl. Movie 4). OHCs were uniquely contractile (Figs. 3a–h, Suppl. Fig. 3), and the temporal dynamics of area changes confirm the reciprocal relationship between electromotility and receptor potential kinetics (Fig. 3a–c; SI, Table 16).
In controls, OHC area changes had $\tau = 60$ ms ($R^2 = 0.866$, Fig. 3a; SI, Table 16); salicylate accelerated this to $\tau = 30$ ms ($R^2 = 0.808$, Fig. 3b); EGTA abolished detectable adaptation (Fig. 3c); and DHS prolonged τ to 190 ms ($R^2 = 0.833$, Fig. 3d). The reticular

lamina (RL) maintained consistent $\tau = 100$ ms ($R^2 = 0.842$ in controls, 0.833 with salicylate; SI, Table 16).

Across five stimulation phases and seven cell/structure types, area, length, and width changes reveal distinct OHC response patterns under different pharmacological conditions (Figs. 3e–h; Suppl. Fig. 3; SI, Tables 17–28). For example, at 105 dB SPL in controls, OHC area reduced by –159 ± 31 µm² ($p = 0.001$, Fig. 3e; SI, Tables 25–28); with salicylate, it reduced further (–200 ± 44 µm², $p = 0.001$, Fig. 3f); with EGTA, contraction was abolished (–4.5 ± 30 µm², $p = 0.500$, Fig. 3g); and with DHS, reduction was attenuated (–89 ± 42 µm², $p = 0.026$, Fig. 3h). Supporting structures behaved differently: TM and BCs expanded significantly (TM: 325 ± 89 µm²; BCs: 227 ± 16 µm², $p = 0.001$), RL expanded moderately (124 ± 18 µm², $p = 0.001$), and Hensen's cells (HeCs) showed variable responses—with notable expansion under salicylate (416 ± 145 µm², $p = 0.040$).

Linear dimensional analyses (Suppl. Figs. 3a–h and 3a–h; SI, Tables 17–24) corroborate these findings. OHC length contraction was significant (–3.5 ± 1.2 µm, $p = 0.018$), enhanced to –5.9 ± 1.1 µm with salicylate ($p = 0.001$; Fig. 1a–d; SI, Tables 17–20). HeCs and RL exhibited opposite-polarity length changes (HeCs: 12.3 ± 3.2 µm, $p = 0.001$; RL: –6.1 ± 1.4 µm, $p = 0.001$), while width changes were less pronounced—mostly expansion, except for OHCs (control: –2.6 ± 0.65 µm, $p = 0.001$; salicylate: –2.5 ± 1.1 µm, $p = 0.040$; Suppl. Fig. 3e–h; SI, Tables 21–24).

A linear mixed-model (LMM) analysis over the 126–183 s exposure (Fig. 3i–k; SI, Tables 29–91) revealed structure-specific morphological changes. For instance, the TM length increased significantly with EGTA (5.39 ± 0.23 µm, $t = 23.63$, $p < 0.001$) but changed minimally with salicylate (0.31 ± 0.19 µm, $t = 1.63$, $p = 0.10$) or DHS (0.11 ± 0.20 µm, $t = 0.55$, $p = 0.58$). Similarly, width changes were non-significant under EGTA (0.20 ± 0.12 µm, $t = 1.69$, $p = 0.09$) and mildly contractile with salicylate (–0.27 ± 0.10 µm, $t = –2.80$, $p < 0.01$). Overall, area expansion under EGTA was significant (264.48 ± 20.99 µm², $t = 12.60$, $p < 0.001$), with minimal DHS (14.93 ± 18.51 µm², $t = 0.81$, $p = 0.42$) and salicylate (–29.56 ± 17.26 µm², $t = –1.71$, $p = 0.09$) effects.

In border cells (BC), EGTA induced significant area expansion (112.07 ± 8.19 µm², $t = 13.68$, $p < 0.001$), with coordinated length (3.32 ± 0.15 µm, $t = 21.96$, $p < 0.001$) and width increases (0.89 ± 0.13 µm, $t = 6.72$, $p < 0.001$). Conversely, salicylate caused area contraction (–38.05 ± 6.73 µm², $t = –5.65$, $p < 0.001$), length reduction (–1.31 ± 0.12 µm, $t = –10.54$, $p < 0.001$), and negligible width change (–0.04 ± 0.11 µm, $t = –0.37$, $p = 0.71$). OHCs responded in a treatment-dependent manner: EGTA increased length (5.58 ± 0.28 µm, $t = 19.91$, $p < 0.001$), produced a mild width increase (0.81 ± 0.17 µm, $t = 4.88$, $p < 0.001$), and expanded area (153.52 ± 7.54 µm², $t = 20.36$, $p < 0.001$); salicylate reduced length (–1.22 ± 0.23 µm, $t = –5.32$, $p < 0.001$), width (–1.08 ± 0.14 µm, $t = –7.90$, $p < 0.001$), and area (–78.07 ± 6.20 µm², $t = –12.59$, $p < 0.001$). HeCs exhibited bidirectional responses: DHS increased length (7.03 ± 0.64 µm, $t = 10.97$, $p < 0.001$) while EGTA caused contraction (–6.96 ± 0.73 µm, $t = –9.58$, $p < 0.001$), with corresponding area changes (DHS: 21.65 ± 19.93 µm², $t = 1.09$, $p = 0.28$; EGTA: –

166.96 ± 22.60 μm², t = –7.39, p < 0.001). The RL elongated with EGTA (2.24 ± 0.31 μm, t = 7.14, p < 0.001) and DHS (1.94 ± 0.28 μm, t = 7.04, p < 0.001) but contracted with salicylate (–1.90 ± 0.26 μm, t = –7.37, p < 0.001). PCs showed robust EGTA-induced elongation (2.27 ± 0.12 μm, t = 18.81, p < 0.001) and area expansion (79.83 ± 3.94 μm², t = 20.25, p < 0.001).

Overall, our results reveal a tight coupling between OHC electromotility and receptor potential kinetics. Under control conditions, while the OHC mechanical response is lowpass-filtered, this leads to sharpened receptor potential kinetics. In contrast, blocking electromotility with salicylate accelerates OHC mechanical response kinetics while producing sluggish, lowpass-filtered receptor potentials. Similarly, reducing extracellular Ca^{2+} with EGTA (100 μM applied at the stereocilia side) abolishes the OHC mechanical response and adaptation—enhancing tectorial membrane expansion and highlighting calcium's dual role in hair cell function and TM integrity[47,51]—while also slowing the electrical response. Conversely, a partial block of the MET channels with extracellular DHS (5 μM applied at the stereocilia side) largely preserves the control-like relationship between electrical and mechanical responses, albeit with a modest delay in mechanical kinetics. Notably, DHS prolongs OHC responses (τ = 190 ms) while maintaining receptor potential sharpness (τ = 20 ms).

This dynamic extends across the cochlear partition, with coordinated deformations of OHCs, HeCs, and the RL during intense stimulation[64,65], consistent with salicylate-induced OHC stiffness decreases[66].

Sound-Evoked Displacements Confirm Electromotility as the Primary Driver of Receptor Potential Kinetic Sharpening

To investigate the relationship between sound-evoked displacements and receptor potential dynamics, we quantified OHC axial and radial displacements concurrent with electrical responses across multiple experimental conditions (Figure 4, see Supplementary Movie 5 for an animated demonstration that synchronizes changes in OHC axial and radial displacements with the electrical response and segmentation visualization).

Under control conditions, a first-order exponential fit of the averaged data from the 126–183 s window during the intensive stimulation phase revealed that outer hair cells (OHCs) exhibited robust axial displacement of 5.02 μm toward the basilar membrane in response to intense acoustic stimulation (105 dB SPL). The displacement kinetics were characterized by a time constant (τ) of 70 ms (R^2 = 0.952, Fig. 4a, see detailed analysis in SI Table 92). In contrast, radial displacements remained below the quantification threshold, precluding reliable exponential fitting (Fig. 4a, SI Table 92). The temporal dynamics of mechanical responses demonstrated pronounced low-pass filtering compared to electrical responses, which exhibited faster kinetics (τ = 10 ms, as described above; Fig. 4a, SI Table 92).

Pharmacological manipulations revealed distinct regulation of mechanical versus electrical adaptation. Salicylate treatment resulted in a modest reduction in axial

displacement amplitude (4.46 μm) while preserving adaptation kinetics (τ = 70 ms, R^2 = 0.921, Fig. 4b, SI Table 92). EGTA exposure led to further amplitude reduction (4.19 μm) with slightly slower adaptation (τ = 110 ms, R^2 = 0.943, Fig. 4c, SI Table 92). Similarly, DHS treatment diminished axial displacement (3.17 μm) while maintaining comparable temporal dynamics (τ = 80 ms, R^2 = 0.967, Fig. 4d, SI Table 92). Across all conditions, radial displacements remained below reliable fitting thresholds, consistent with the predominantly axial nature of OHC motility (SI Table 92).

Strikingly, superimposition with the electrical recordings did not exhibit the reciprocal relationship described above between electromotility and electrical recordings, confirming the specificity of OHC electromotility, not sound-evoked displacements, as the cochlear amplifier (Figs 2-3 vs Fig. 4).

Comprehensive displacement analyses across the 7 cell/structure types (Fig. 4e–l, SI Table 93-108) under sound stimulation revealed complex mechanical coupling that extends well beyond outer hair cell OHC) motion. Under control conditions (Fig. 4e, Table 93-108), OHCs showed robust axial displacement (−5.62 ± 0.59 μm at 105 dB SPL), but outer pillar cell heads (PCs), reticular lamina (RL), inner hair cell/inner pillar cell complex (IHC), border cells (BC), and the tectorial membrane (TM) also moved in a coordinated yet distinct manner. Notably, Hensen's cells (HeCs) exhibited the largest axial responses (−7.70 ± 1.13 μm, $p < 0.001$ vs. baseline), while the TM displayed the smallest displacement (−1.65 ± 0.34 μm, $p < 0.001$ vs. baseline). This descending hierarchy of motion (HeCs > PC > RL > IHC > BC > TM) supports prior observations that the organ's pivot points reside near the modiolus[67]. Radial displacements were consistently smaller than axial ones in all structures, although HeCs again exhibited the largest radial movements in controls (−5.44 ± 0.79 μm at 105 dB SPL, $p < 0.001$ vs. other structures, Fig. 4i, SI Table 93-108).

Pharmacological interventions that decouple electrical and mechanical responses demonstrated pronounced, structure-specific effects (Fig. 4f–l, SI Table 93-108). Salicylate application (Fig. 4f, j, Fig. 4e–l, SI Table 93-108) significantly reduced OHC displacement (from −5.62 to −4.45 μm, $p < 0.001$) but paradoxically enhanced HeC axial motion (−8.98 ± 1.27 μm, $p < 0.001$ vs. baseline), suggesting intricate mechanical coupling between these cell types. In contrast, EGTA treatment (Fig. 4g, k, Fig. 4e–l, SI Table 93-108), which abolishes electromotility, preserved substantial OHC displacement (−4.72 ± 0.36 μm, $p < 0.001$) but markedly reduced HeC motion (−3.84 ± 2.18 μm, $p < 0.001$). Notably, EGTA maintained significant displacement in PCs (−4.72 ± 0.49 μm, $p < 0.001$). DHS exposure (Fig. 4h, l, SI Table 93-108) also showed strong, cell-specific outcomes, most prominently affecting HeCs (−8.98 ± 2.19 μm, $p < 0.001$ vs. baseline) while retaining moderate OHC movement (−3.94 ± 0.66 μm, $p < 0.001$). In all pharmacological conditions, radial responses remained smaller than axial ones, although HeCs consistently displayed relatively large radial displacements compared to other structures (Fig. 4e–l, SI Table 93-108).

Linear mixed-model analysis across 126–183 s window (thereby effectively evaluating the displacement time constants) revealed strong, structure-specific treatment effects

(Fig. 4m-n, SI Table 109-136). In the axial direction (Fig. 4m, SI Table 109-136), HeCs exhibited the most pronounced changes ($F = 84.39$, $p < 0.001$), showing enhancement under DHS (2.14 ± 0.29 µm, $t = 7.51$, $p < 0.001$) and salicylate (1.71 ± 0.27 µm, $t = 6.41$, $p < 0.001$), but suppression under EGTA (-2.59 ± 0.32 µm, $t = -8.01$, $p < 0.001$). OHCs also demonstrated strong treatment sensitivity ($F = 31.35$, $p < 0.001$), with both DHS (-0.61 ± 0.09 µm, $t = -6.97$, $p < 0.001$) and salicylate (-0.69 ± 0.08 µm, $t = -8.47$, $p < 0.001$) causing significant reductions. In the radial direction (Fig. 4n, SI Table 109-136), HeCs again showed the largest responses ($F = 344.52$, $p < 0.001$), highlighted by a pronounced expansion under DHS (5.18 ± 0.22 µm, $t = 23.17$, $p < 0.001$). Overall, this quantitative analysis confirms the reciprocal relationship between OHCs and HeCs described above, whose responses shifted in opposing directions across treatments—most notably under salicylate, where HeC enhancement (axial: 1.71 ± 0.27 µm, radial: 1.91 ± 0.21 µm) contrasted sharply with OHC suppression (axial: -0.69 ± 0.08 µm, radial: -0.13 ± 0.10 µm).

These findings reinforce the reciprocal OHC–HeC relationship identified above in the morphological change analysis, wherein HeCs expand in opposition to OHC contraction. Although this coordinated behavior magnifies overall cochlear partition motion, it does not contribute to receptor potential sharpening. Instead, the temporal precision of electrical responses remains unique to OHC electromotility, underscoring that sound-evoked displacements alone cannot account for the rapid kinetics of the receptor potential. Collectively, these results confirm a robust mechanical feedback network that leverages HeC mechanics to enhance OHC-driven forces, while firmly establishing OHC electromotility as the fundamental driver of sharper receptor potential kinetics.

Cycle-by-Cycle Analysis Reveals OHC Area Changes During Mild Sound Stimulation

To validate area changes as a metric for electromotility under physiologically relevant sound intensity, we developed a high-resolution imaging approach for detecting subtle mechanical responses to mild acoustic stimulation. This required overcoming two key technical challenges: the spatial resolution limitations of whole-organ imaging and the need for reliable structure identification in reflection mode. We addressed these by implementing a targeted imaging strategy focused on the first-row OHC apical region, which provides distinct anatomical landmarks for segmentation while maintaining structure identity (Figure 5a-c). Super-resolution imaging confirmed this region encompasses the OHC apical surface from near its center to just below the hair bundle within the reticular lamina (Figure 5d-f).

We integrated time-resolved confocal microscopy with precise auditory stimulation using a pixel clock-based master timing system (Figure 5g). This setup enabled synchronized image acquisition with sound delivery, allowing temporal phase analysis across 12 equal segments of the stimulus cycle[68]. AI-based segmentation (Figure 5h-j, IoU = 0.87, Pixel Accuracy = 0.95; see Suppl. Fig. 4 for more performance characterization vs the 7 class-segmentation above) was then applied to quantify mechanical metrics on a cycle-by-cycle basis.

At 70 dB SPL and 250 Hz stimulation—a 56-fold reduction in sound pressure compared to our earlier protocol—OHCs exhibited consistent cyclic motions. Peak-to-peak amplitudes reached 98.69 ± 19.94 nm axially and 29.75 ± 21.38 nm radially (n = 13 OHCs from independent preparations; Figure 5k-m). These nanoscale movements align with previous measurements using optical coherence tomography and interferometry in the guinea pig cochlear apex[69]. Notably, OHCs showed cyclic area reductions of 2.96 ± 0.86 µm² from a baseline of 109.22 ± 1.66 µm² (2.7% change), consistent with established OHC length changes of 3-10%[3,5,6,70]. These findings demonstrate that area changes provide a reliable metric for quantifying electromotility at physiologically relevant sound intensities.

Discussion

These findings resolve a longstanding paradox in cochlear mechanics: how outer hair cells (OHCs) maintain precise high-frequency tuning despite the nominal low-pass filtering imposed by their membrane resistance–capacitance (RC) properties. Intuitively, OHC voltage responses should roll off at frequencies well below the upper limit of mammalian hearing, leading some to propose that hair-bundle motility—rather than somatic electromotility—dominates amplification at higher frequencies. However, by examining the relatively slow, DC-like mechanical responses of OHCs (on tens-to-hundreds of milliseconds timescales) in direct conjunction with their electrical signals, we show that this slow mechanical actuator *sharpens* the receptor potential rather than attenuating it.

A key demonstration of this principle is the dramatic reversal caused by salicylate, which blocks prestin: OHCs then undergo *faster* mechanical changes, yet the receptor potentials become *sluggish* and low-pass filtered. By contrast, under control conditions, OHC electromechanics are slower (tens of milliseconds), but the receptor potential remains crisp (on the order of milliseconds or faster). These results indicate that the nominal "RC limit" does not compromise high-frequency hearing; instead, low-pass mechanical filtering in the OHC body selectively dampens slower voltage distortions, effectively preserving the sharper, higher-frequency components in the receptor potential.

Although our data come from the apical region (best frequency ~180 Hz)—an area chosen for optimal confocal imaging—the principle uncovered here should extend throughout the cochlea. High-frequency OCT and interferometry studies have shown OHC somatic motility at tens of kilohertz[40], implying that this mechanical–electrical interplay applies across a wide spectral range. Indeed, recent reports demonstrate that OHCs can exhibit robust electromechanical effects well above 10 kHz[43], reinforcing the generalizability of our findings. The apex thus serves as an experimentally accessible window onto universal OHC mechanisms.

Moreover, these experiments highlight the distinction between DC (or slow) and AC (rapid) components of OHC activity. While hair-bundle adaptation can operate on sub-millisecond scales, we focused on slower DC electromotility that becomes prominent at higher sound intensities. Far from undercutting fast transduction, these slower mechanical processes *filter out* detrimental low-frequency voltage changes, thereby maintaining the

fast portion of the receptor potential. This clarifies why OHC electromotility and high-frequency hearing remain compatible: the "slow" mechanical response shapes the voltage domain in a way that preserves rapid signals.

These findings also illuminate interactions with supporting cells such as Hensen's cells (HeCs). Salicylate treatments, for example, diminished OHC motion while paradoxically enhancing HeC displacement. Although this helps magnify overall partition movements, it does not reintroduce the fast receptor potential that requires intact prestin-driven electromotility. This further confirms that OHC somatic activity is the primary driver of fine-tuned, high-frequency signal amplification, whereas passive supporting-cell expansions modulate the overall micromechanics. Although salicylate is typically ototoxic at high doses[71], our findings suggest it may enable mechanical and electrical recovery after intense stimulation under controlled conditions. This raises the possibility that low-level salicylate or related compounds could provide therapeutic strategies for noise-induced hearing loss.

Methodologically, the AI-powered segmentation and 2D measurements of OHC area changes proved more sensitive than single-axis (length or width) analyses. Even under moderate acoustic stimulation, we resolved submicrometer changes that correlate well with known electromotile amplitudes[3,5,6,70]. Future studies should extend high-resolution, cycle-by-cycle imaging into higher-frequency or basal-cochlear regions. This would more directly link the present DC measurements with the microsecond-scale adaptation events at kHz rates, bridging DC electromotility and high-frequency cochlear mechanics in real time.

From an evolutionary standpoint, low-pass OHC mechanics may be precisely what enables *high-frequency* electrical fidelity in mammals[38,72–75], distinguishing them from non-mammalian vertebrates whose ears lack prestin-based electromotility. Rather than acting as a bottleneck, this "slow" actuator proves essential for preserving rapid transduction components, conferring exquisite frequency selectivity. By showing how DC electromotility and receptor potentials reinforce each other, our study reveals a broader principle of mechanical–electrical co-filtering in sensory cells. Ultimately, these data underscore that the OHC's unique electromechanical properties are not just compatible with, but *vital* for, the unparalleled tuning and sensitivity of mammalian hearing.

Materials and Methods

Detailed experimental procedures are provided in the Supplementary Methods. Briefly, guinea pigs (2-5 weeks old, either sex) were handled according to approved ethical guidelines (DNR 5111-2019, DNR 01122-2020 ID2529, DNR 19058-2021 ID4016), with procedures performed as previously described[45,47,48,51] and recently updated[45]. After temporal bone isolation and cochlear exposure, two openings were made for perfusion and recording, with middle ear immersion attenuating sound stimuli by approximately 25 dB SPL. Glass microelectrodes (impedance ~3 MΩ) filled with artificial endolymph were inserted through Reissner's membrane under confocal imaging guidance, targeting a region with best frequency around 180 Hz (see Figure 1 of ref[45] for electrode location),

while electrical signals were amplified and digitized at 10 kHz. The experimental setup synchronized acoustic stimulation (tone bursts at 160 Hz, Suppl. Fig. 1), electrical recording, and confocal imaging using a customized LabVIEW interface, with imaging performed on a Carl Zeiss LSM 980 confocal microscope equipped for external triggering.

Acknowledgments
P.H. acknowledges the grants FB23-0003 FB24-0005 from the Swedish Tysta skolan foundation and M3 from the Torsten Söderberg foundation. H.S.P. acknowledges multiple grants from the Research Fund for the Center for Laboratory Medicine, St. Olavs hospital. Dr. Mikael Rousson is thanked for initial discussions on cochlear image segmentation strategies.

Data Availability Statement
Statement: All data supporting the findings of this study are available within the manuscript and its Supplementary Information files. Raw datasets are available from the corresponding author upon reasonable request, data transfer agreements and associated costs may be required.

References
1. Dallos, P. The active cochlea. *J. Neurosci.* **12**, 4575–4585 (1992).
2. Liberman, M. C. *et al.* Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature* **419**, 300–304 (2002).
3. Brownell, W. E., Bader, C. R., Bertrand, D. & de Ribaupierre, Y. Evoked mechanical responses of isolated cochlear outer hair cells. *Science* **227**, 194–196 (1985).
4. Zheng, J. *et al.* Prestin is the motor protein of cochlear outer hair cells. *Nature* **405**, 149–155 (2000).
5. Ashmore, J. Cochlear outer hair cell motility. *Physiol. Rev.* **88**, 173–210 (2008).
6. Ashmore, J. F. A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J. Physiol.* **388**, 323–347 (1987).
7. Fettiplace, R. & Kim, K. X. The physiology of mechanoelectrical transduction channels in hearing. *Physiol. Rev.* **94**, 951–986 (2014).
8. Eatock, R. A., Corey, D. P. & Hudspeth, A. J. Adaptation of mechanoelectrical transduction in hair cells of the bullfrog's sacculus. *J. Neurosci. Off. J. Soc. Neurosci.* **7**, 2821–2836 (1987).

9. Howard, J. & Hudspeth, A. J. Mechanical relaxation of the hair bundle mediates adaptation in mechanoelectrical transduction by the bullfrog's saccular hair cell. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 3064–3068 (1987).

10. Assad, J. A. & Corey, D. P. An active motor model for adaptation by vertebrate hair cells. *J. Neurosci. Off. J. Soc. Neurosci.* **12**, 3291–3309 (1992).

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

12. Holt, J. R. & Corey, D. P. Two mechanisms for transducer adaptation in vertebrate hair cells. *Proc. Natl. Acad. Sci.* **97**, 11730–11735 (2000).

13. Stauffer, E. A. & Holt, J. R. Sensory transduction and adaptation in inner and outer hair cells of the mouse auditory system. *J. Neurophysiol.* **98**, 3360–3369 (2007).

14. Peng, A. W., Effertz, T. & Ricci, A. J. Adaptation of mammalian auditory hair cell mechanotransduction is independent of calcium entry. *Neuron* **80**, 960–972 (2013).

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

16. Assad, J. A., Shepherd, G. M. & Corey, D. P. Tip-link integrity and mechanical transduction in vertebrate hair cells. *Neuron* **7**, 985–994 (1991).

17. Hudspeth, A. J. How the ear's works work. *Nature* **341**, 397–404 (1989).

18. Pickles, J. O., Comis, S. D. & Osborne, M. P. Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. *Hear. Res.* **15**, 103–112 (1984).

19. Beurg, M., Fettiplace, R., Nam, J.-H. & Ricci, A. J. Localization of inner hair cell mechanotransducer channels using high-speed calcium imaging. *Nat. Neurosci.* **12**, 553–558 (2009).

20. Kachar, B., Parakkal, M., Kurc, M., Zhao, Y. & Gillespie, P. G. High-resolution structure of hair-cell tip links. *Proc. Natl. Acad. Sci.* **97**, 13336–13341 (2000).

21. Schwander, M., Kachar, B. & Müller, U. Review series: The cell biology of hearing. *J. Cell Biol.* **190**, 9–20 (2010).

22. Holton, T. & Hudspeth, A. J. The transduction channel of hair cells from the bull-frog characterized by noise analysis. *J. Physiol.* **375**, 195–227 (1986).

23. Crawford, A. C., Evans, M. G. & Fettiplace, R. Activation and adaptation of transducer currents in turtle hair cells. *J. Physiol.* **419**, 405–434 (1989).

24. Tan, X., Beurg, M., Hackney, C., Mahendrasingam, S. & Fettiplace, R. Electrical tuning and transduction in short hair cells of the chicken auditory papilla. *J. Neurophysiol.* **109**, 2007–2020 (2013).

25. Beurg, M., Evans, M. G., Hackney, C. M. & Fettiplace, R. A large-conductance calcium-selective mechanotransducer channel in mammalian cochlear hair cells. *J. Neurosci. Off. J. Soc. Neurosci.* **26**, 10992–11000 (2006).

26. Géléoc, G. S., Lennan, G. W., Richardson, G. P. & Kros, C. J. A quantitative comparison of mechanoelectrical transduction in vestibular and auditory hair cells of neonatal mice. *Proc. Biol. Sci.* **264**, 611–621 (1997).

27. He, D. Z. Z., Jia, S. & Dallos, P. Mechanoelectrical transduction of adult outer hair cells studied in a gerbil hemicochlea. *Nature* **429**, 766–770 (2004).

28. Jia, S., Dallos, P. & He, D. Z. Z. Mechanoelectric transduction of adult inner hair cells. *J. Neurosci. Off. J. Soc. Neurosci.* **27**, 1006–1014 (2007).

29. Kros, C. J., Rusch, A. & Richardson, G. P. Mechano-Electrical Transducer Currents in Hair Cells of the Cultured Neonatal Mouse Cochlea. *Proc. R. Soc. Lond. Ser. B* **249**, 185–193 (1992).

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

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

32. Dallos, P., Santos-Sacchi, J. & Flock, A. Intracellular recordings from cochlear outer hair cells. *Science* **218**, 582–584 (1982).

33. Kennedy, H. J., Evans, M. G., Crawford, A. C. & Fettiplace, R. Fast adaptation of mechanoelectrical transducer channels in mammalian cochlear hair cells. *Nat. Neurosci.* **6**, 832–836 (2003).

34. Ruggero, M. A. & Rich, N. C. Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane. *J. Neurosci.* **11**, 1057–1067 (1991).

35. Evans, B. N. & Dallos, P. Stereocilia displacement induced somatic motility of cochlear outer hair cells. *Proc. Natl. Acad. Sci.* **90**, 8347–8351 (1993).

36. Nuttall, A. L. & Ren, T. Electromotile hearing: evidence from basilar membrane motion and otoacoustic emissions. *Hear. Res.* **92**, 170–177 (1995).

37. Fridberger, A. & de Monvel, J. B. Sound-induced differential motion within the hearing organ. *Nat. Neurosci.* **6**, 446–448 (2003).

38. Fettiplace, R. Diverse mechanisms of sound frequency discrimination in the vertebrate cochlea. *Trends Neurosci.* **43**, 88 (2020).

39. Fettiplace, R. & Hackney, C. M. The sensory and motor roles of auditory hair cells. *Nat. Rev. Neurosci.* **7**, 19–29 (2006).

40. Dewey, J. B., Altoè, A., Shera, C. A., Applegate, B. E. & Oghalai, J. S. Cochlear outer hair cell electromotility enhances organ of Corti motion on a cycle-by-cycle basis at high frequencies in vivo. *Proc. Natl. Acad. Sci.* **118**, e2025206118 (2021).

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

42. Martin, P. & Hudspeth, A. J. Active hair-bundle movements can amplify a hair cell's response to oscillatory mechanical stimuli. *Proc. Natl. Acad. Sci.* **96**, 14306–14311 (1999).

43. Olson, E. S. & Strimbu, C. E. Cochlear mechanics: new insights from vibrometry and optical coherence tomography. *Curr. Opin. Physiol.* **18**, 56–62 (2020).

44. Gao, J. *et al.* Prestin-based outer hair cell electromotility in knockin mice does not appear to adjust the operating point of a cilia-based amplifier. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 12542–12547 (2007).

45. Hakizimana, P. The summating potential polarity encodes the ear health condition. *Cell. Mol. Life Sci. CMLS* **80**, 163 (2023).

46. Caprara, G. A., Mecca, A. A. & Peng, A. W. Decades-old model of slow adaptation in sensory hair cells is not supported in mammals. *Sci. Adv.* **6**, eabb4922 (2020).

47. Hakizimana, P. & Fridberger, A. Inner hair cell stereocilia are embedded in the tectorial membrane. *Nat. Commun.* **12**, 2604 (2021).

48. Ulfendahl, M., Flock, A. & Khanna, S. M. Isolated cochlea preparation for the study of cellular vibrations and motility. *Acta Oto-Laryngol. Suppl.* **467**, 91–96 (1989).

49. Hakizimana, P., Brownell, W. E., Jacob, S. & Fridberger, A. Sound-induced length changes in outer hair cell stereocilia. *Nat. Commun.* **3**, 1094 (2012).

50. Hakizimana, P. The sensitivity of mechanoelectrical transduction response phase to acoustic overstimulation is calcium-dependent. *Pflugers Arch.* **476**, 271–282 (2024).

51. Strimbu, C. E., Prasad, S., Hakizimana, P. & Fridberger, A. Control of hearing sensitivity by tectorial membrane calcium. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 5756–5764 (2019).

52. Fridberger, A., Tomo, I., Ulfendahl, M. & Boutet de Monvel, J. Imaging hair cell transduction at the speed of sound: dynamic behavior of mammalian stereocilia. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 1918–1923 (2006).

53. Pettersen, H. S. *et al.* Code-Free Development and Deployment of Deep Segmentation Models for Digital Pathology. *Front. Med.* **8**, 816281 (2021).

54. Santos-Sacchi, J., Kakehata, S. & Takahashi, S. Effects of membrane potential on the voltage dependence of motility-related charge in outer hair cells of the guinea-pig. *J. Physiol.* **510**, 225–235 (1998).

55. Ricci, A. Differences in Mechano-Transducer Channel Kinetics Underlie Tonotopic Distribution of Fast Adaptation in Auditory Hair Cells. *J. Neurophysiol.* **87**, 1738–1748 (2002).

56. Corns, L. F. *et al.* Mechanotransduction is required for establishing and maintaining mature inner hair cells and regulating efferent innervation. *Nat. Commun.* **9**, 4015 (2018).

57. Brownell, W. E., Jacob, S., Hakizimana, P., Ulfendahl, M. & Fridberger, A. Membrane cholesterol modulates cochlear electromechanics. *Pflugers Arch.* **461**, 677–686 (2011).

58. Ricci, A. J., Kennedy, H. J., Crawford, A. C. & Fettiplace, R. The Transduction Channel Filter in Auditory Hair Cells. *J. Neurosci.* **25**, 7831–7839 (2005).

59. Ricci, A. J., Crawford, A. C. & Fettiplace, R. Active Hair Bundle Motion Linked to Fast Transducer Adaptation in Auditory Hair Cells. *J. Neurosci.* **20**, 7131–7142 (2000).

60. Ricci, A. J., Wu, Y.-C. & Fettiplace, R. The Endogenous Calcium Buffer and the Time Course of Transducer Adaptation in Auditory Hair Cells. *J. Neurosci.* **18**, 8261–8277 (1998).

61. Flock, A., Bretscher, A. & Weber, K. Immunohistochemical localization of several cytoskeletal proteins in inner ear sensory and supporting cells. *Hear. Res.* **7**, 75–89 (1982).

62. Gueta, R., Barlam, D., Shneck, R. Z. & Rousso, I. Measurement of the mechanical properties of isolated tectorial membrane using atomic force microscopy. *Proc. Natl. Acad. Sci.* **103**, 14790–14795 (2006).

63. Ghaffari, R., Page, S. L., Farrahi, S., Sellon, J. B. & Freeman, D. M. Electrokinetic properties of the mammalian tectorial membrane. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4279–4284 (2013).

64. Flock, A., Flock, B., Fridberger, A., Scarfone, E. & Ulfendahl, M. Supporting cells contribute to control of hearing sensitivity. *J. Neurosci. Off. J. Soc. Neurosci.* **19**, 4498–4507 (1999).

65. Fridberger, A., Flock, A., Ulfendahl, M. & Flock, B. Acoustic overstimulation increases outer hair cell Ca2+ concentrations and causes dynamic contractions of the hearing organ. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7127–7132 (1998).

66. Kakehata, S. & Santos-Sacchi, J. Effects of Salicylate and Lanthanides on Outer Hair Cell Motility and Associated Gating Charge. *J. Neurosci.* **16**, 4881–4889 (1996).

67. Fridberger, A., Monvel, J. B. de & Ulfendahl, M. Internal Shearing within the Hearing Organ Evoked by Basilar Membrane Motion. *J. Neurosci.* **22**, 9850–9857 (2002).

68. Jacob, S., Tomo, I., Fridberger, A., de Monvel, J. B. & Ulfendahl, M. Rapid confocal imaging for measuring sound-induced motion of the hearing organ in the apical region. *J. Biomed. Opt.* **12**, 021005 (2007).

69. Burwood, G., Hakizimana, P., Nuttall, A. L. & Fridberger, A. Best frequencies and temporal delays are similar across the low-frequency regions of the guinea pig cochlea. *Sci. Adv.* **8**, eabq2773.

70. Holley, M. C. & Ashmore, J. F. On the mechanism of a high-frequency force generator in outer hair cells isolated from the guinea pig cochlea. *Proc. R. Soc. Lond. B Biol. Sci.* **232**, 413–429 (1988).

71. Sheppard, A., Hayes, S., Chen, G.-D., Ralli, M. & Salvi, R. Review of salicylate-induced hearing loss, neurotoxicity, tinnitus and neuropathophysiology. *Acta Otorhinolaryngol. Ital.* **34**, 79 (2014).

72. Sasmal, A. & Grosh, K. Unified cochlear model for low- and high-frequency mammalian hearing. *Proc. Natl. Acad. Sci.* **116**, 13983–13988 (2019).

73. Meaud, J. & Grosh, K. Response to a Pure Tone in a Nonlinear Mechanical-Electrical-Acoustical Model of the Cochlea. *Biophys. J.* **102**, 1237–1246 (2012).

74. Ashmore, J. F. Frequency tuning in a frog vestibular organ. *Nature* **304**, 536–538 (1983).

75. Hill, E. M., Koay, G., Heffner, R. S. & Heffner, H. E. Audiogram of the chicken (Gallus gallus domesticus) from 2 Hz to 9 kHz. *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* **200**, 863–870 (2014).

76. Hakizimana, P. & Fridberger, A. Effects of salicylate on sound-evoked outer hair cell stereocilia deflections. *Pflugers Arch.* **467**, 2021–2029 (2015).

77. Boutet de Monvel, J., Le Calvez, S. & Ulfendahl, M. Image Restoration for Confocal Microscopy: Improving the Limits of Deconvolution, with Application to the Visualization of the Mammalian Hearing Organ. *Biophys. J.* **80**, 2455–2470 (2001).

78. Tassano, M., Delon, J. & Veit, T. FastDVDnet: Towards Real-Time Deep Video Denoising Without Flow Estimation. in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 1351–1360 (2020). doi:10.1109/CVPR42600.2020.00143.

Figures and Tables

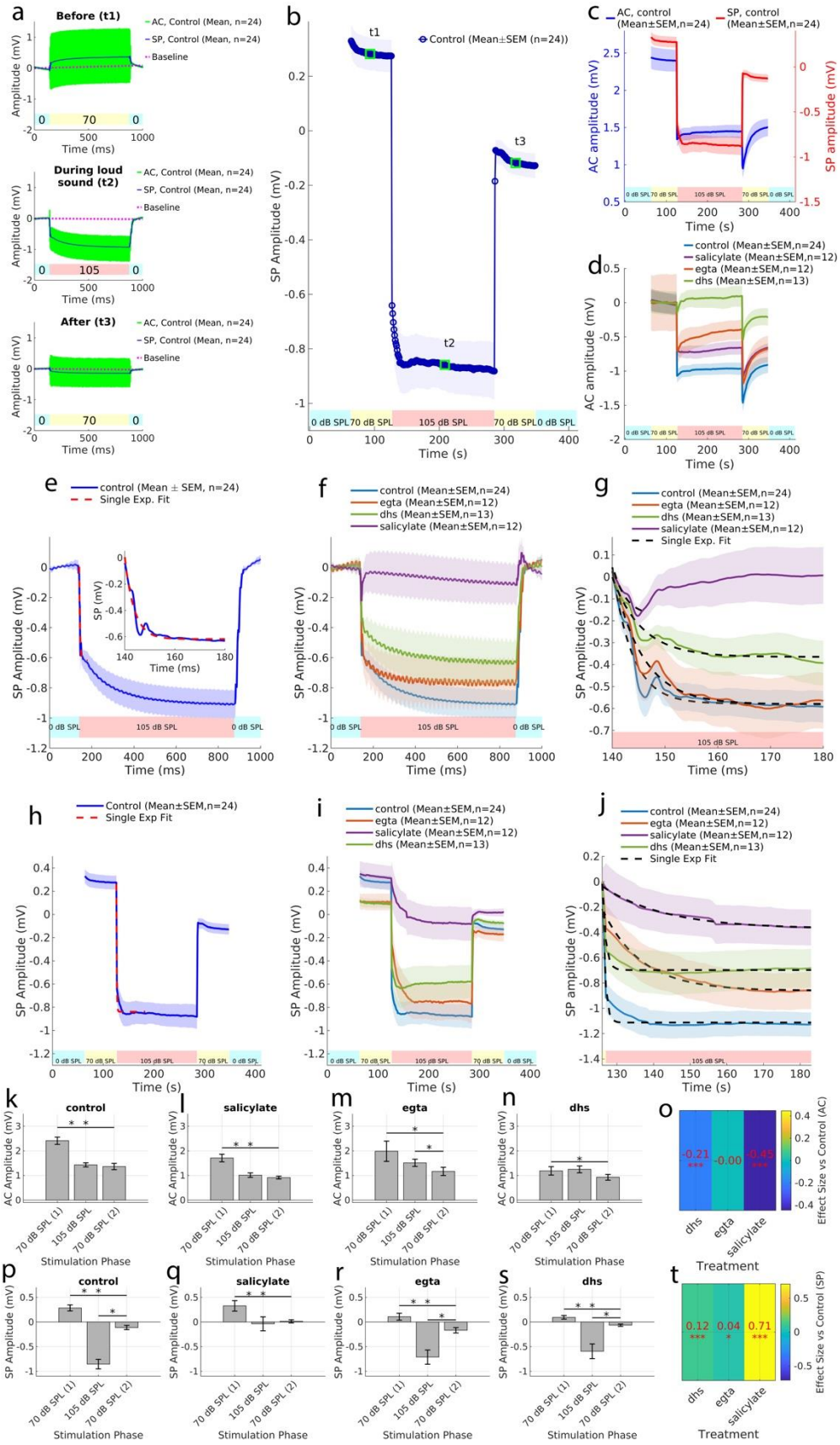

a
Before (t1)
AC, Control (Mean, n=24)
SP, Control (Mean, n=24)
Baseline
Amplitude (mV)
Time (ms)
During loud sound (t2)
After (t3)
b
Control (Mean±SEM (n=24))
t1
t2
t3
SP Amplitude (mV)
Time (s)
0 dB SPL
70 dB SPL
105 dB SPL
c
AC, control (Mean±SEM,n=24)
SP, control (Mean±SEM,n=24)
AC amplitude (mV)
SP amplitude (mV)
d
control (Mean±SEM,n=24)
salicylate (Mean±SEM,n=12)
egta (Mean±SEM,n=12)
dhs (Mean±SEM,n=13)
e
control (Mean ± SEM, n=24)
Single Exp. Fit
SP (mV)
f
g
h
Control (Mean±SEM,n=24)
Single Exp Fit
i
j
k
control
l
salicylate
m
egta
n
dhs
AC Amplitude (mV)
70 dB SPL (1)
105 dB SPL
70 dB SPL (2)
Stimulation Phase
o
-0.21
-0.00
-0.45
Effect Size vs Control (AC)
Treatment
p
q
r
s
SP Amplitude (mV)
t
0.12
0.04
0.71
Effect Size vs Control (SP)

Figure 1. Prestin-Mediated Electromotility Orchestrates Multi-Timescale Cochlear Adaptation

High-temporal resolution recordings show receptor potential adaptation patterns in mammalian cochlea. (a) Electrical responses showing AC (green) and SP (blue) components during stepped acoustic stimulation (70 → 105 → 70 dB SPL) at t_1=93s, t_2=209s, and t_3=318s (n=24). (b) Biphasic SP adaptation showing rapid decrease followed by plateau during high-intensity stimulation. (c) Dynamic coupling between AC and SP amplitudes throughout stimulation. (d) AC responses under pharmacological interventions: control, prestin-blocking salicylate, calcium-chelating EGTA, and MET channel-blocking DHS.

Fast adaptation analysis (e-g): (e) Control response with exponential fit showing 60% adaptation within 140-180 ms. (f) Treatment-specific profiles revealing distinct salicylate pattern: adaptation completed within 0.85 stimulus cycles, followed by recovery and plateau. (g) Comparative kinetics analyzed within specified windows for all conditions. Salicylate exhibited slowest adaptation (τ = 0.46 ms).

Slow adaptation analysis (h-j): (h) Control recordings with exponential fit over 57s window. (i) SP amplitude traces for all treatments showing distinct adaptation profiles. (j) Exponential fits demonstrate EGTA prolongs kinetics while reducing amplitude, DHS preserves kinetics but lowers amplitude, and salicylate reduces amplitude with slowest kinetics.

Quantitative analysis (k-t): AC amplitudes for all conditions (k-n) with LMM analysis (o). SP amplitudes (p-s) with LMM analysis of adaptation time constants (t). Conditions: Control (n=24), salicylate (10 mM, n=12), EGTA (100 µM, n=12), and DHS (5µM, n=13). *$p < 0.05$, **$p < 0.01$. Acoustic stimuli: 160 Hz, sampled at 10,000 Hz.

Colored bars indicate stimulus levels (cyan: 0 dB SPL, yellow: 70 dB SPL, red: 105 dB SPL).

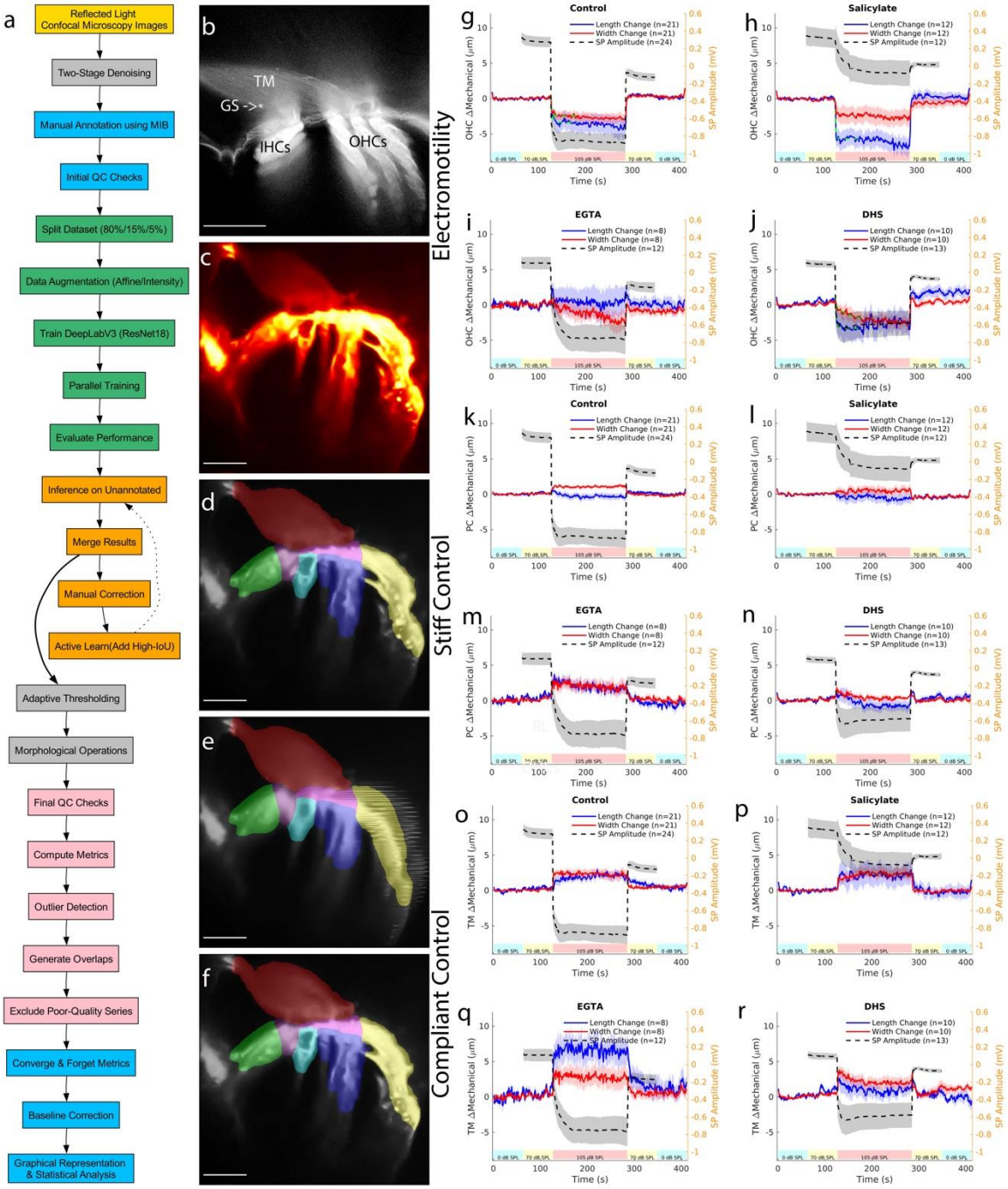


Figure 2. Outer Hair Cell Electromotility Sharpens Receptor Potential Kinetics
(a) Segmentation pipeline: sequential processing of reflected light confocal microscopy images through denoising, annotation, deep learning segmentation, tracking, and validation. (b) Fluorescence microscopy showing labeled cochlear structures: tectorial membrane (TM), granular structures (GS[47]), inner hair cells (IHCs), and outer hair cells (OHCs). (c) Reflected light confocal microscopy of same region showing enhanced structural detail. (d-f) Segmented images of cochlear structures before, during, and after loud sound stimulation. Color code: brown (TM), green (border cells), dark magenta

(IHCs/inner pillar complex), cyan (outer pillar cell heads (PCs)), clear magenta (reticular lamina), blue (OHCs), yellow (Hensen's cells). See Supplementary Movies 1-2.

(g-j) OHC electromotility responses compared with SP amplitude data from Figure 1: (g) Control conditions: 105 dB SPL induced length changes (-3.53 ± 1.9 µm) and width changes (-2.81 ± 0.6 µm, n=21), time constants 60 ms and 50 ms respectively. (h) Salicylate accelerated kinetics (τ = 20 ms for length). (i) EGTA attenuated responses below reliable fitting threshold. (j) DHS showed differential effects: moderate length kinetics (τ = 50 ms) versus slower width changes (τ = 450 ms). See Supplementary Movie 3.

(k-n) Pillar cell responses: No detectable adaptation across all conditions (control, salicylate, EGTA, DHS).

(o-r) Tectorial membrane responses: Consistent 2-3 µm expansion in both dimensions across conditions. EGTA enhanced expansion, aligned with TM's calcium-dependent properties. Scale bars: 50 µm.

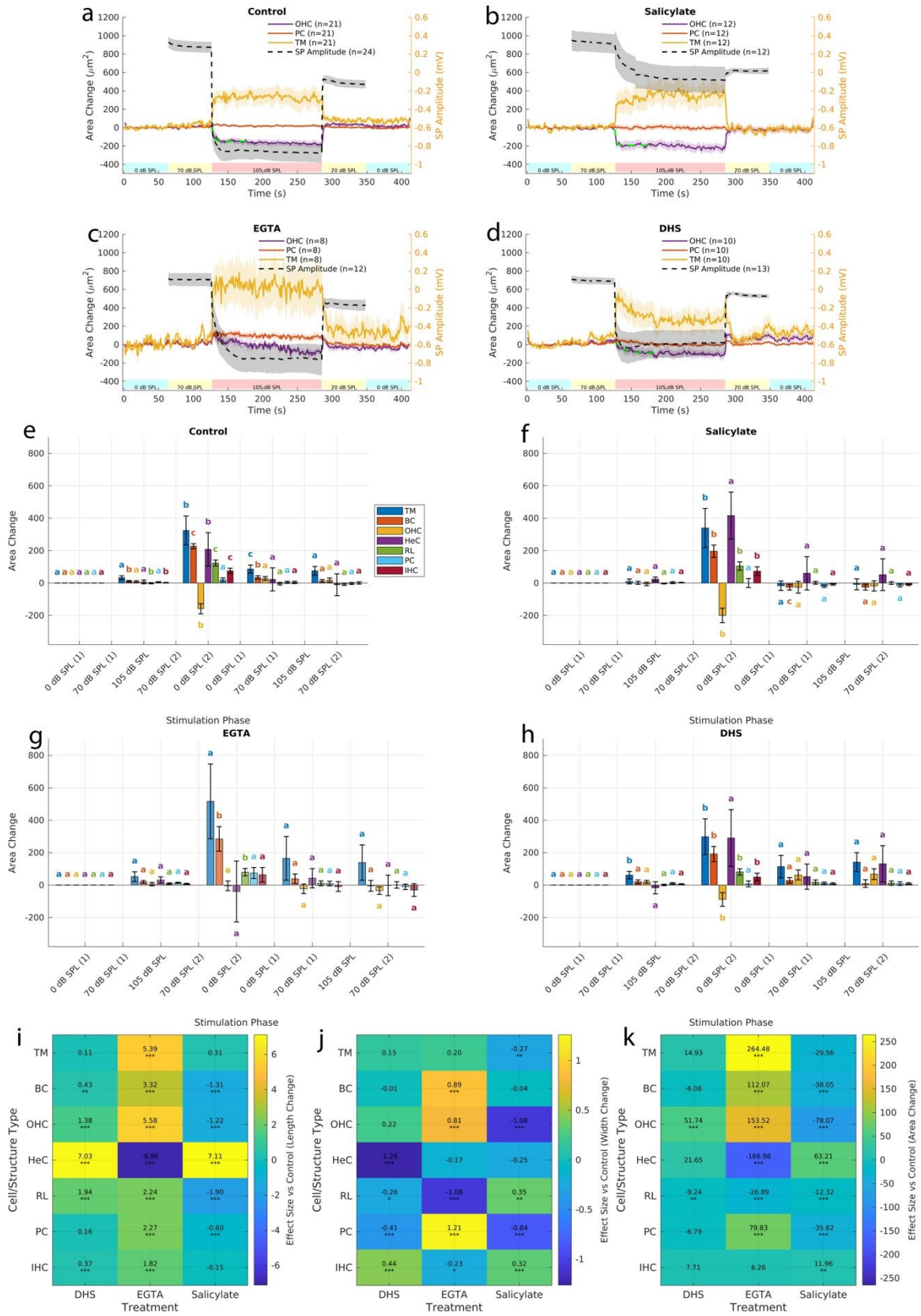

a
Control
OHC (n=21)
PC (n=21)
TM (n=21)
SP Amplitude (n=24)
b
Salicylate
OHC (n=12)
PC (n=12)
TM (n=12)
SP Amplitude (n=12)
c
EGTA
OHC (n=8)
PC (n=8)
TM (n=8)
SP Amplitude (n=12)
d
DHS
OHC (n=10)
PC (n=10)
TM (n=10)
SP Amplitude (n=13)
Area Change (μm²)
SP Amplitude (mV)
Time (s)
0 dB SPL
70 dB SPL
105 dB SPL
e
Control
f
Salicylate
g
EGTA
h
DHS
Area Change
TM
BC
OHC
HeC
RL
PC
IHC
0 dB SPL (1)
70 dB SPL (1)
105 dB SPL
70 dB SPL (2)
0 dB SPL (2)
Stimulation Phase
i
j
k
Cell/Structure Type
Treatment
DHS
EGTA
Salicylate
Effect Size vs Control (Length Change)
Effect Size vs Control (Width Change)
Effect Size vs Control (Area Change)

Figure 3. Area change confirms electromotility role in sharpening receptor potential kinetics.

(a-d) Area changes and receptor potential during acoustic stimulation under different conditions. Control OHC showed characteristic time constant, salicylate accelerated response, DHS prolonged it, while EGTA abolished mechanical adaptation. SP data from Figure 1 included for timescale comparison. See Supplementary Movie 4.

(e-h) Area changes across stimulation phases under different conditions. Results shown using color-coded compact letter display (shared letters: $p \geq 0.05$, different letters: $p < 0.05$). Control: OHC area reduction during 105 dB SPL, with TM and BC expansion. Salicylate enhanced OHC contraction. EGTA abolished OHC responses. DHS attenuated contraction. Hensen's cells showed pronounced expansion under salicylate.

(i-k) Linear mixed-model analysis during intense sound exposure (126-183 s): (i) Length changes: EGTA induced elongation in TM and OHCs. (j) Width changes: notable OHC contraction under salicylate. (k) Area changes: EGTA induced expansion in TM and OHCs; salicylate induced OHC contraction. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

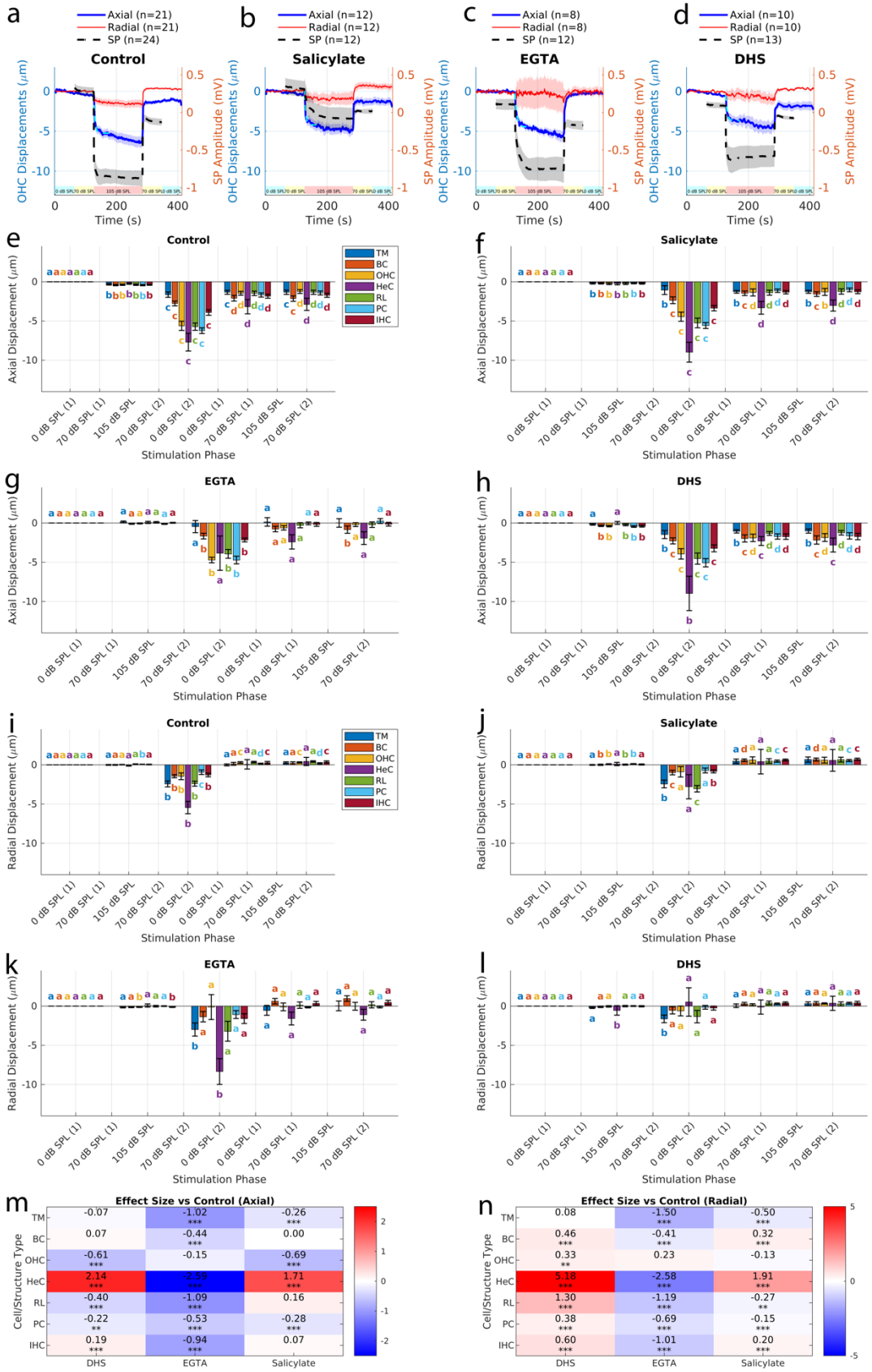

a
Axial (n=21)
Radial (n=21)
SP (n=24)
Control
b
Axial (n=12)
Radial (n=12)
SP (n=12)
Salicylate
c
Axial (n=8)
Radial (n=8)
SP (n=12)
EGTA
d
Axial (n=10)
Radial (n=10)
SP (n=13)
DHS
OHC Displacements (μm)
SP Amplitude (mV)
Time (s)
e
f
g
h
i
j
k
l
Axial Displacement (μm)
Radial Displacement (μm)
Stimulation Phase
0 dB SPL (1)
70 dB SPL (1)
105 dB SPL
70 dB SPL (2)
0 dB SPL (2)
TM
BC
OHC
HeC
RL
PC
IHC
m
Effect Size vs Control (Axial)
n
Effect Size vs Control (Radial)
Cell/Structure Type
DHS
EGTA
Salicylate

Figure 4. Sound-evoked displacements confirm electromotility as the primary driver of receptor potential sharpening.

(a-d) Temporal dynamics of OHC axial and radial displacements with concurrent SP recordings. Control conditions showed robust axial displacement while radial displacements remained below quantification threshold. SP data showed faster kinetics. Salicylate reduced axial displacement while maintaining time constant. EGTA reduced displacement amplitude and slowed adaptation. DHS diminished axial displacement while preserving temporal dynamics. See Supplementary Movie 5.

(e-h) Axial displacement patterns across cochlear structures during stimulation phases. Results shown using color-coded compact letter display (shared letters: $p \geq 0.05$, different letters: $p < 0.05$). Control conditions showed largest displacement in HeCs and smallest in TM. Salicylate reduced OHC displacement while enhancing HeC motion. EGTA preserved OHC displacement but reduced HeC motion. DHS maintained moderate OHC movement with enhanced HeC response.

(i-l) Radial displacement patterns showed consistently smaller movements compared to axial displacements. HeCs exhibited the largest radial movements across all conditions, particularly pronounced under control conditions.

(m-n) Linear mixed-model analysis of treatment effects on (m) axial and (n) radial displacements during intense sound exposure. HeCs showed strongest treatment sensitivity, with opposing responses to OHCs, particularly under salicylate. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

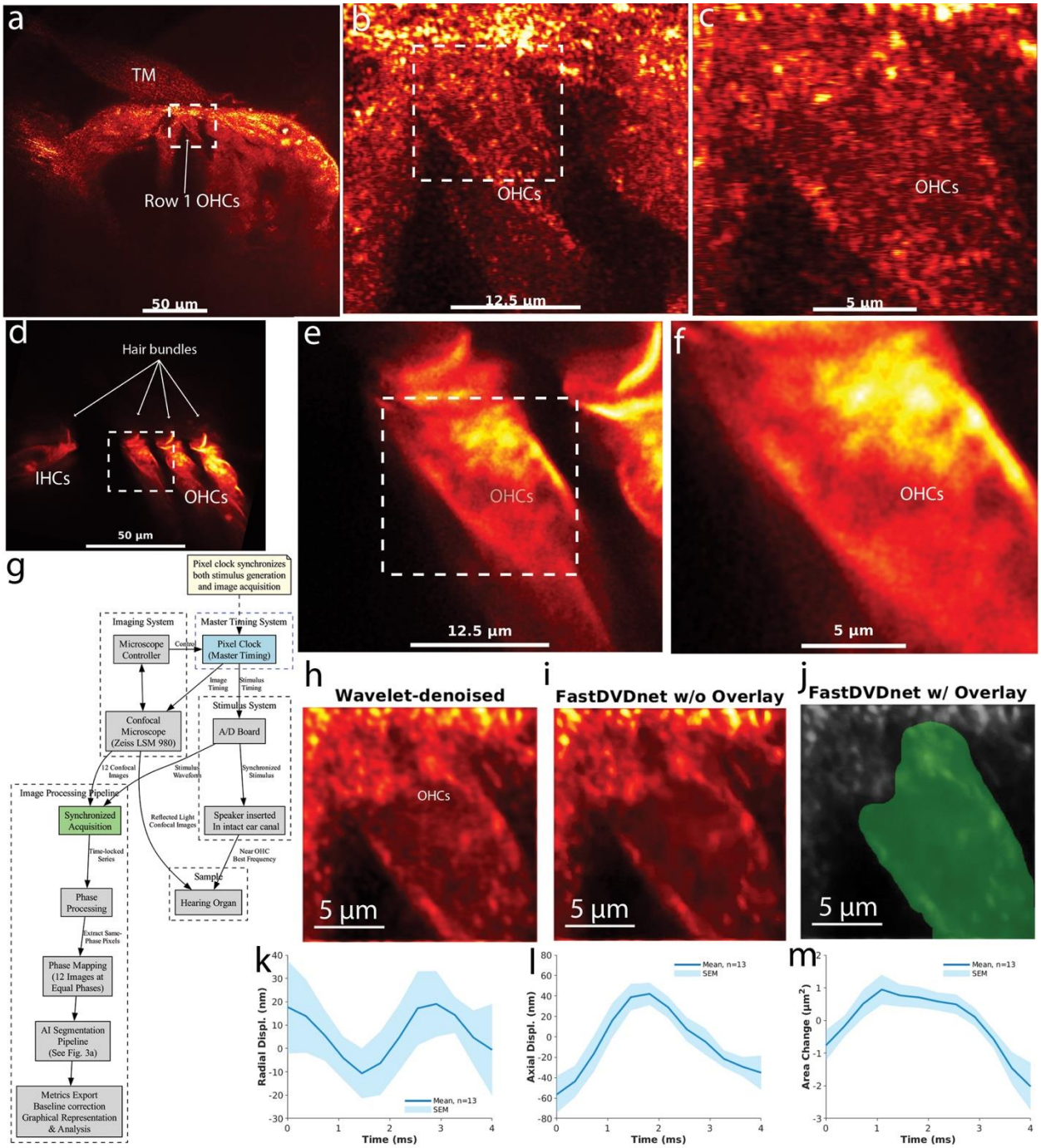


Figure 5. Cycle-by-Cycle OHC Displacements and Area Changes During Mild Sound Stimulation. (a-c) Reflection-mode confocal images of first-row OHC apical region at increasing magnification, showing anatomical landmarks between reticular lamina, outer pillar cell head, and row 2 OHC for reliable tracking.

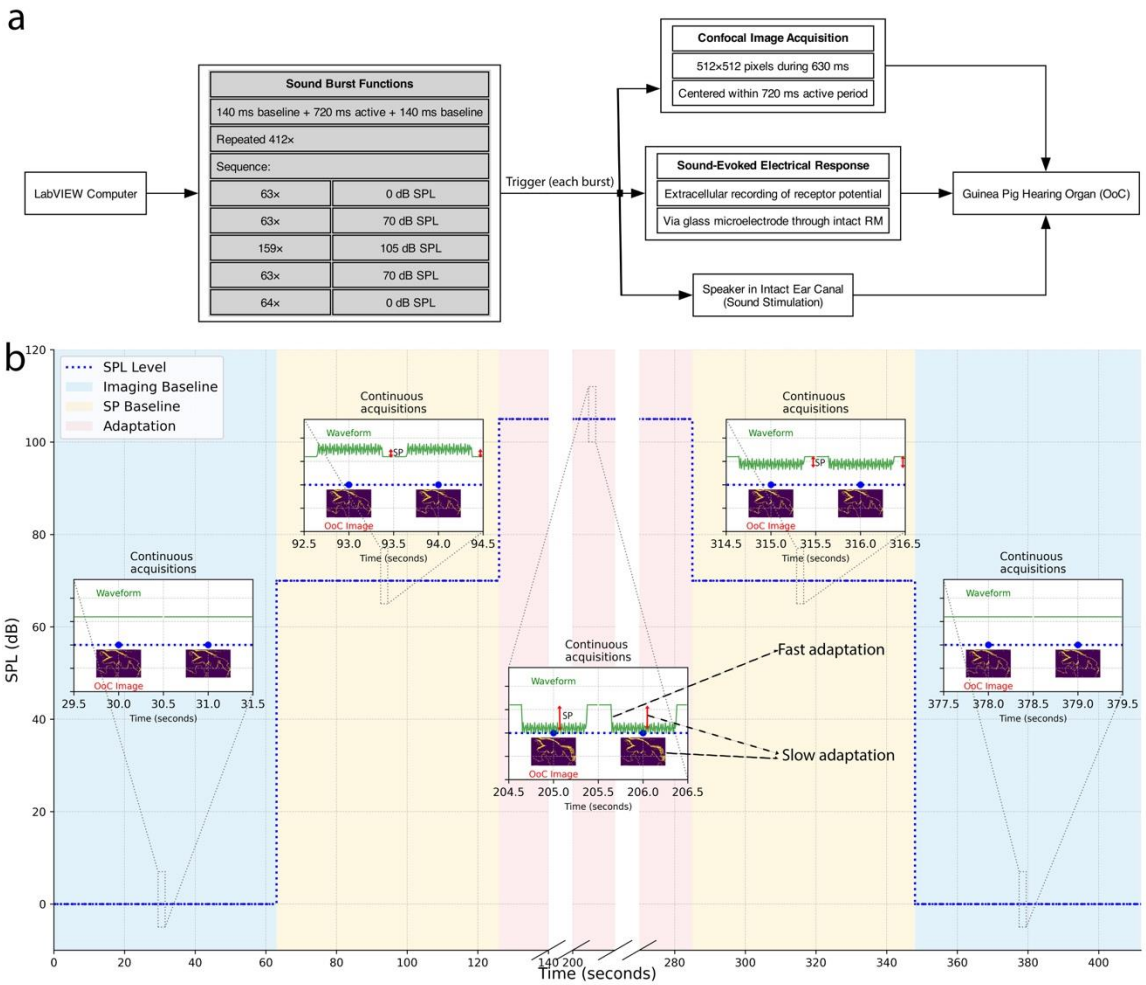
(d-f) Super-resolution images confirming field of view captures OHC surface from near-center to sub-bundle region, encompassing the apical half of the hair cell.

(g) Integrated system schematic showing confocal microscope synchronized with acoustic stimulation. System combines pixel clock timing, LSM 980 imaging, and stimulus delivery for phase-mapped analysis of OHC responses.

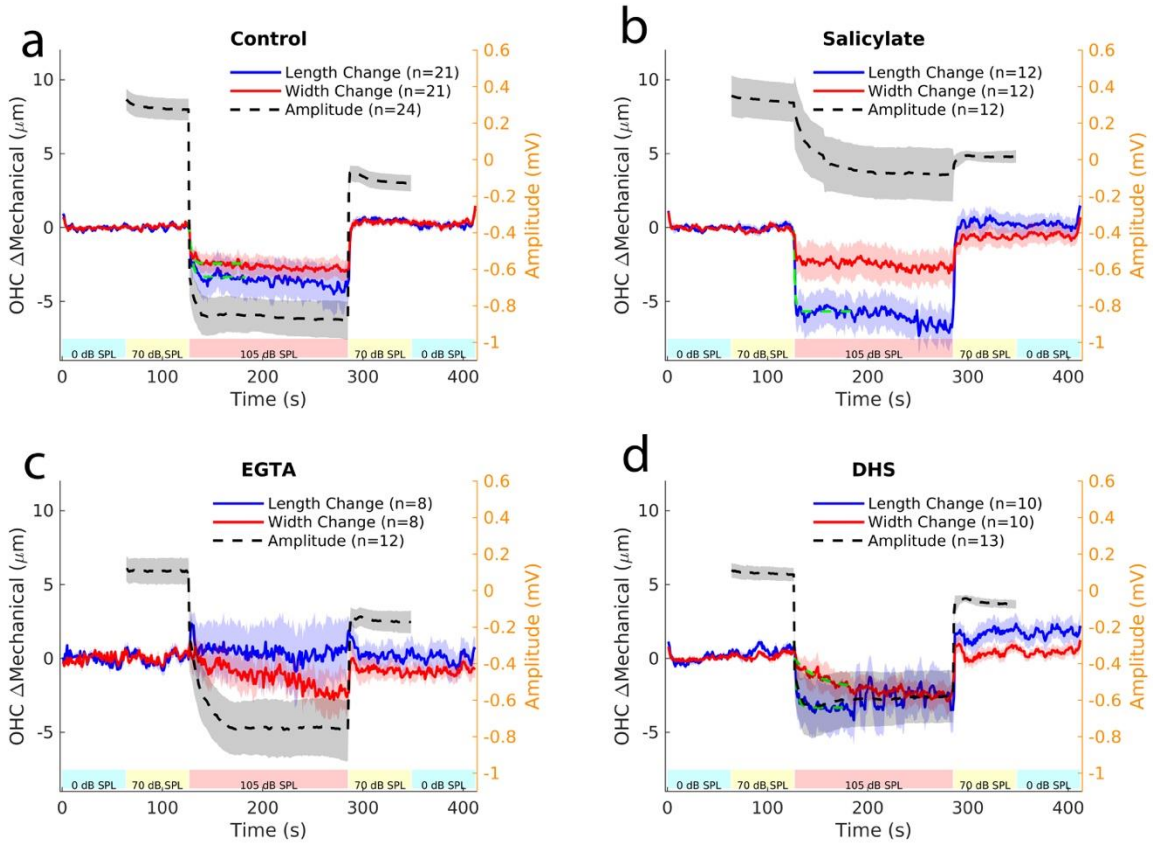
(h-j) Image processing sequence: wavelet denoising, FastDVDnet processing, and AI-based segmentation of row 1 OHC apical half region.

(k-m) Time-resolved analysis during mild sound stimulation showing radial displacement, axial displacement, and surface area changes, averaged across 13 cells from different preparations.

Supplementary Figures

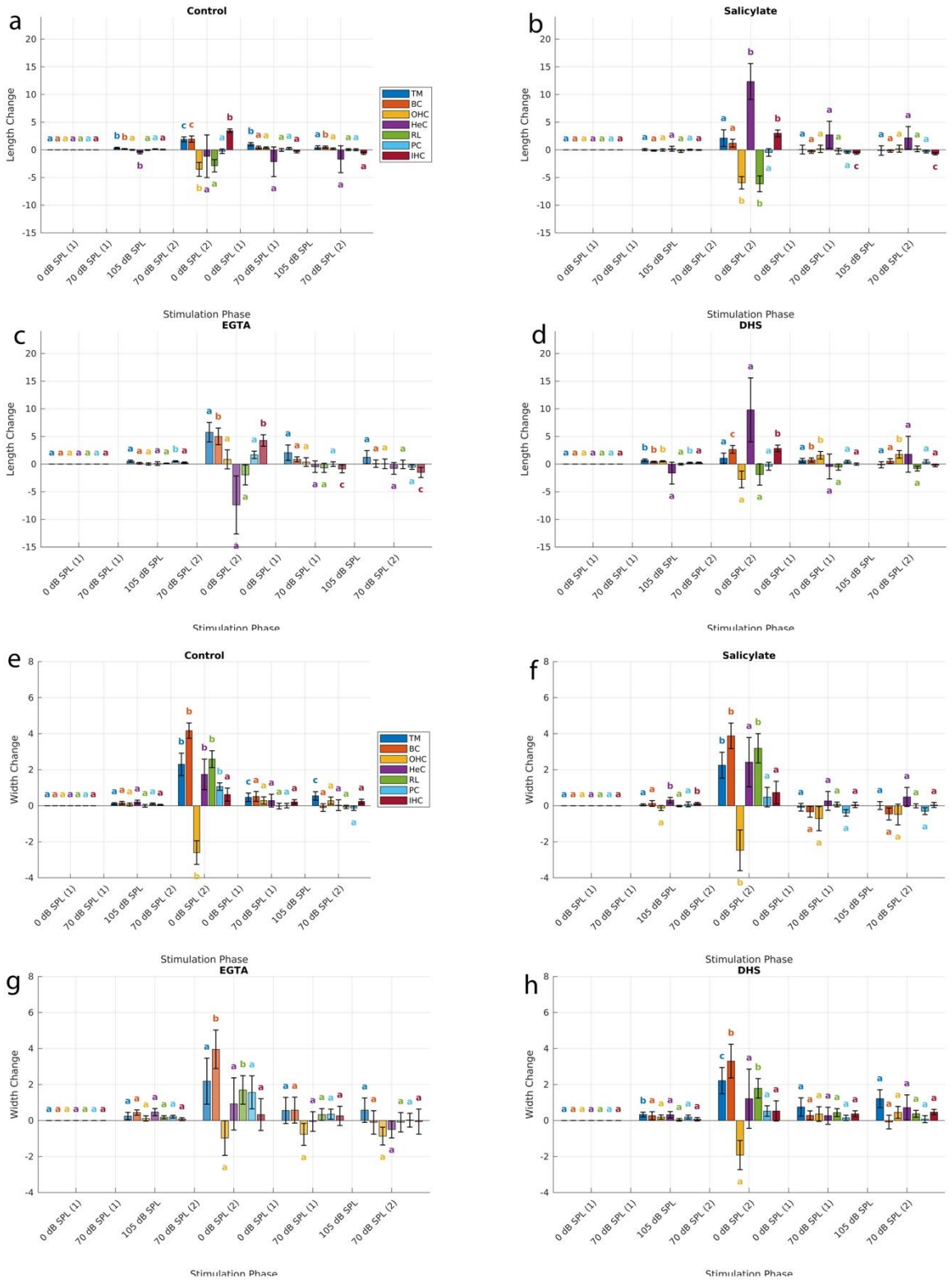


Supplementary Figure 1. Experimental design for monitoring synchronous mechanical and electrical responses in the guinea pig hearing organ during sound exposure. A) Schematic of the synchronized measurement system. Each measurement cycle acquires electrical signals and confocal images during a 1000-millisecond period: the sound burst (720 ms) is flanked by two 140-millisecond quiet intervals, with microscope image acquisition aligned to the central 630 ms of sound exposure. Quiet intervals provide baseline reference for electrical response amplitude calculations. (B) Complete experimental timeline illustrating sound intensity variations (blue dotted line) and measurements over 412 seconds. The protocol comprises five sequential phases at different sound intensities: initial baseline (0 dB SPL (1)), initial moderate sound (70 dB SPL (1)), sustained loud sound (105 dB), recovery period (70 dB SPL (2)), and final baseline (0 dB SPL (2)). Insets illustrate electrical recordings (green waveforms) alongside organ structure outlines from microscopy images (edge detection, OpenCV). Background colors indicate experimental phases: blue (silence), yellow (moderate sound), red (loud sound).



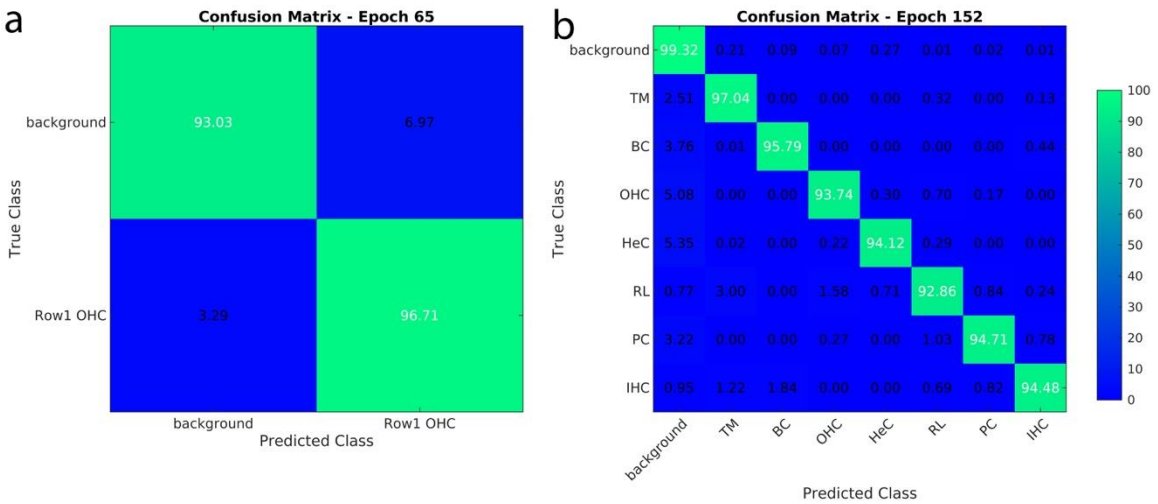
Supplementary Figure 2. Outer Hair Cell Electromotility Sharpens Receptor Potential Kinetics.

Panels (a–d) display the same data as Fig. 2g–j, but with magnified views to enhance visualization of the exponential fits (green dotted lines). See detailed analysis in SI Table 15. Colored bars along the abscissa indicate stimulus level transitions (cyan: 0 dB SPL, yellow: 70 dB SPL, red: 105 dB SPL).



Supplementary Figure 3. Quantification of Length and Width Changes across Five Stimulation Phases in Seven Cell/Structure Types Reveals Distinct Patterns in OHC Responses under Various Pharmacological Conditions. For clarity, the typical

horizontal bar with p-value stars format is replaced by a color-coded compact letter display (CLD): bars sharing the same letter are not significantly different ($p \geq 0.05$), whereas bars lacking a common letter are significantly different ($p < 0.05$). Panels (a–d) show length change measurements in control (a), salicylate (b), EGTA (c), and DHS (d). In control conditions, OHCs exhibited significant contraction during intense stimulation (–3.5 ± 1.2 μm, $p = 0.018$), and further contracted in presence of salicylate (–5.9 ± 1.1 μm, $p = 0.001$). Notably, Hensen's cells (HeCs) and OHCs exhibited a reciprocal relationshipt; see SI Tables 17–20). Panels (e–h) present width change measurements in control (e) and in the presence of salicylate (f), EGTA (g), and DHS (h). Width changes were generally less pronounced and more uniform across structures, with most exhibiting expansion, except for OHCs, which contracted under control (–2.6 ± 0.65 μm, $p = 0.001$) and salicylate (–2.5 ± 1.1 μm, $p = 0.040$) conditions. Detailed analysis is provided in SI Tables 21–24. Abbreviations: TM, tectorial membrane; BCs, border cells; IHCs/IPC, inner hair cells/inner pillar cell complex; PCs, outer pillar cell heads; RL, reticular lamina; OHCs, outer hair cells; HeCs, Hensen's cells.



Supplementary Figure 4. Performance of the DeepLabV3-Resnet18 deep segmentation models for cochlear structure recognition.
(a) Normalized confusion matrix showing the classification outcomes for the single outer hair cell (OHC) segmentation, where the two classes are background and Row1 OHC.
(b) Normalized confusion matrix depicting the classification outcomes for eight cochlear structures: Exterior; TM (tectorial membrane); BC (border cells); OHC (outer hair cells); HeC (Hensen's cells); RL (reticular lamina); PCs (outer pillar cell heads); and IHCs (inner hair cells/inner pillar cell complex). The matrices compares the predicted classes (x-axis) with the actual classes (y-axis) on a test set of fully annotated time series that were not used during network training. Green colors represent a higher proportion of instances of the ground truth class being classified as the predicted class.

Legends For Supplementary Videos

Supplementary Movie 1: Time-series acquired by reflected light confocal imaging (see Methods) after 2D-wavelet and FastDVDnet denoising. Total sequence of 412 frames acquired during: baseline (first 63 frames, 0 dB SPL), mild stimulation (next 63 frames, 70 dB SPL), loud stimulation (next 159 frames, 105 dB SPL), recovery (next 63 frames, 70 dB SPL), and extended recovery (final 64 frames, 0 dB SPL).

Supplementary Movie 2: Same as in Supplementary Movie 1 after AI-powered segmentation of cochlear structures using DeepLab V3 plus Resnet18. The segmentation highlights key structures: tectorial membrane (brown), border cells (green), IHCs/inner pillar cell complex (dark magenta), outer pillar cell head (cyan), reticular lamina (clear magenta), outer hair cells (blue), and Hensen's cells (yellow). An animated demonstration of the segmentation visualization.

Supplementary Movie 3: Animated demonstration that synchronizes changes in OHC length and width with its electrical response and segmentation visualization.

Supplementary Movie 4: Animated demonstration that synchronizes changes in OHC, PC, and TM areas with the electrical response and segmentation visualization

Supplementary Movie 5: Animated demonstration that synchronizes changes in OHC axial and radial displacements with the electrical response and segmentation visualization.

Supplementary Material and Methods

Guinea Pig Temporal Bone Preparation

Guinea pigs (2–5 weeks old, either sex) were used in this study, conducted per ethical guidelines approved by the Regional Ethics Committee in Linköping (DNR 5111-2019, DNR 01122-2020 ID2529, DNR 19058-2021 ID4016).

Animal handling and surgical procedures were performed as previously described[45,47,48,51] and updated recently[45]. Briefly, guinea pigs were anesthetized and decapitated; the temporal bone was isolated, the bulla opened, and the middle and inner ear immersed in oxygenated minimum essential medium with Earle's balanced salts. Under a stereomicroscope, the cochlea was exposed and two small openings were made: one at the base to connect a perfusion tube with oxygenated medium (extending cochlear viability up to 4 hours[57]), and one at the apex for perfusion exit, confocal imaging, and recording sound-evoked electrical responses.

Middle ear immersion attenuated sound stimuli by approximately 25 dB SPL, which was accounted for in reported intensities.

A rotatable holder optimized the apical angle for cross-sectional visualization with reflected light confocal microscopy; the chamber was immobilized under the LSM 980 confocal microscope within a light-isolated Faraday cage on a vibration-dampening table. For the picture of the preparation, see ref[57].

Electrode Preparation and Insertion

Glass microelectrodes (1.5 mm outer diameter) were prepared[45,47,48,51] using a standard puller and beveled at a 20° angle to achieve an impedance of approximately 3 MΩ, corresponding to a tip diameter of ~10 µm. The electrodes were filled with artificial endolymph, composed of 1.3 mM NaCl, 31 mM $KHCO_3$, 128.3 mM KCl, and 0.023 mM $CaCl_2$ (pH 7.4), with osmolality adjusted to 300 mOs/kg using sucrose.

The beveled electrodes were mounted on a computer-controlled micromanipulator equipped with a step motor. Under confocal imaging guidance, the electrode was inserted through Reissner's membrane (RM) and parallel to the stria vascularis. Perpendicular penetration of the RM and near the stria vascularis ensured smooth, controlled entry, an approach facilitated by the cross-sectional orientation. After penetration, microscope imaging parameters were optimized to obtain detailed views of the cochlear structures close to the electrode (see Figure 1 of ref[45] for the electrode location near the hearing organ).

The insertion targeted a region with the best frequency of around 180 Hz. Electrical signals were amplified using an Ix1 or NPI BA-01X amplifier (10× gain) and digitized at 10 kHz using a 24-bit A/D converter (NI USB-4431).

Acoustic Stimulation

Tone bursts of 1000 ms were generated at 160 Hz, slightly below the best frequency (~180 Hz) at the recording location[45,51]. Each burst comprised a 720 ms active sound period flanked by 140 ms silent intervals, facilitating measurement of the summating potential[45,50,51]. To investigate adaptation, the sound intensity was automatically stepped through a preloaded sequence of five distinct levels without any delay between bursts (Suppl. Fig. 1):

- 0 dB SPL for 63 bursts (baseline establishment),
- 70 dB SPL for 63 bursts,
- 105 dB SPL for 159 bursts,
- 70 dB SPL for another 63 bursts, and
- 0 dB SPL for 64 bursts (final baseline).

The entire burst sequence was preprogrammed and loaded into the system before the experiment. Once all preparations were complete, the sequence was initiated with a single mouse click, allowing the program to execute the sequence automatically without requiring further human intervention.

Sound bursts were delivered through a speaker tube stably inserted into the ear canal of the temporal bone preparation, ensuring precise and consistent sound delivery throughout the experiment.

Synchronized Electrophysiology and Confocal Imaging

A key feature of this setup was the precise synchronization of acoustic stimulation, electrical recording, and confocal imaging (Suppl. Fig. 1). A customized LabVIEW (National Instruments) interface delivered the trigger signals that simultaneously:

1. Presented the tone burst to the cochlea,
2. Initiated extracellular receptor potential recordings, and
3. Triggered acquisition of high-speed confocal images.

Imaging was performed on a Carl Zeiss LSM 980 confocal microscope equipped with a macro interface, allowing external triggers from LabVIEW. Once in "pairing mode," the confocal system captured each image during the central portion of the 720 ms active sound period, yielding direct, temporally aligned comparisons between electrophysiological waveforms and structural changes. Over the course of a full

experimental sequence, a total of 412 electrophysiological traces and confocal frames were acquired: 63 at 0 dB SPL (pre-stimulation), 63 at 70 dB SPL, 159 at 105 dB SPL, 63 again at 70 dB SPL, and 64 at 0 dB SPL (post-stimulation). For detailed procedures for imaging, see below.

Pharmacological Administration

Experiments were conducted using fresh temporal bone preparations, including both control and drug-treated conditions. For controls, preparations without any drug treatment established baseline responses. In experimental groups, 100 μM EGTA or 5 μM DHS was dissolved in the endolymph before electrode loading and beveling; drug release was achieved via gentle pressure (2–4 psi). For salicylate treatments, 10 mM salicylate was introduced into the perfusion medium, allowing equilibration for approximately 10 minutes or until around 0.1 mL perfusion volume decrease[49,57,76]. Cochlear microphonics were routinely monitored to confirm preparation stability prior to starting the experiments; in drug-treated preparations, this included monitoring before and after drug administration.

SP Waveforms Computation

Each raw waveform was baseline-corrected using a linear fit from the first and last 1000 points (samples 1–1000 and 9001–10000) using MATLAB's polyfit, applied over the entire trace with polyval. SP amplitudes were calculated as the difference between the mean of the baseline-corrected signal (samples 2000–8500) and the baseline. Values were corrected for 10× amplification and converted to millivolts.

AC and SP Amplitude Computations

AC amplitudes were estimated from the raw electrical waveforms using a moving window approach. Peak-to-peak amplitudes were calculated within windows of size Fsamp/StimFreq (10 kHz/160 Hz), corresponding to the stimulus period. The mean of these local peak-to-peak amplitudes was taken as the AC amplitude estimate, converted to millivolts, and corrected for the 10× amplification. For SP traces in Figure 2, SP components were extracted from the raw electrical waveforms by (1) applying a linear polynomial fit to the start and end points for baseline correction and (2) isolating the SP using a moving mean filter of 63 samples (fsamp = 10,000 Hz; stimulus frequency = 160 Hz). The resulting amplitudes were also corrected for the 10× amplification and converted to millivolts. To facilitate real-time monitoring during acquisition, 90% of the raw electrical waveforms were stored directly in real time, while 10% were captured as a rolling average. These rolling-average points were later interpolated using MATLAB's interp1 function to create continuous datasets for curve fitting and graphical presentation.

First order exponential Curve Fitting

The first-order exponential curve fitting was performed using an automated approach to ensure consistency across electrical and mechanical datasets under various experimental conditions. This methodology adhered to stringent criteria for selecting fitting windows and initializing parameters, thereby maintaining uniformity and reliability throughout the analysis. For slow adaptation processes, specifically in both electrical and mechanical

data, the fitting was applied over a consistent time window of 126–183 s. This range was carefully selected to capture the dynamics of slow adaptation reliably. In mechanical data, the exponential behavior varied depending on structural and experimental conditions, with the trend of the segment—whether increasing or decreasing—dictating the choice of exponential model. The algorithm autonomously determined the appropriate model for each dataset using a decision-making function based on the sign of the slope over the fitting segment: segments with positive slopes were identified as increasing (requiring a growing exponential model), while those with negative slopes were classified as decreasing (requiring a decaying exponential model).

Two distinct exponential models were employed: a decaying model, $y = c + ae^{-bt}$, for processes with decreasing trends, and a growing model, $y = c + a(1 - e^{-bt})$, for processes exhibiting growth. This automated model selection ensured that the fitting procedure was applied consistently across all datasets.

To ensure robust fitting, boundary constraints were applied uniformly across all datasets, regardless of whether they represented electrical, mechanical, or pharmacological conditions. The constraints were carefully chosen to reflect biologically plausible values while maintaining consistency. Lower bounds were set to $[0, 0.1, \min(y)]$, ensuring that the amplitude (a) was non-negative, the rate (b) was not excessively slow, and the offset (c) was above the minimum observed value. Upper bounds were set to $[2 \times range(y), 100, \max(y)]$, preventing unrealistic amplitudes, excessively fast dynamics, or offsets exceeding the observed maximum. This consistent application of boundary constraints ensured that all fits were directly comparable across different electrical and mechanical datasets.

To confirm the quality of the fits, each fitted curve was plotted against the raw data and visually inspected to assess its alignment with the observed trends. Only fits with an R^2 value greater than or equal to 0.8 were deemed reliable and retained for further analysis. Time was normalized to a [0, 1] range for numerical stability and executed the fitting procedure using MATLAB's fit function. The selection of the exponential model, based on whether the trend was growing or decaying, enabled accurate characterization of adaptation processes. For visualization, fitted curves were superimposed on the raw data, with distinct dashed lines representing growth (cyan) and decay (green).
Fit parameters, including amplitude, rate, and offset, were reported alongside goodness-of-fit metric R^2. Additionally, time constants (τ=1/|b|) were calculated to characterize the kinetics of adaptation.

Reflected Light Confocal Imaging
The Zeiss LSM 980 (ZEN blue software, excitation 488 nm) microscope's increased sensitivity allowed operation at lower gain (~500 V). We used a 40×, 0.8 NA Nikon NIR APO water immersion lens at 1× zoom (pixel size 0.52 µm), achieving 512×512 images in 630 ms.

Time-Resolved Confocal Imaging

To capture sound-evoked motion dynamics of the apical row 1 OHC region, we employed the same setup in reflection mode with ZEN blue software, utilizing time-resolved acquisition mode[68]. We used up to 16× zoom, capturing 12 time-synchronized images during exposure to a 250 Hz, 70 dB SPL tone.

Fluorescence Confocal Imaging

Fluorescence imaging (Fig. 3a) used di-3-anepdhq dye in endolymph for staining[47]. Images were acquired on an LSM 780 with 0.8 NA NIR APO water immersion lens (1.4× digital zoom, 0.36 µm/pixel). Settings: 561 nm excitation, 566-689 nm detection, 900 V gain, 31 s acquisition with 4 averages. Despite this extended acquisition time, fewer cochlear structures were visible compared to reflected light imaging (630 ms). The super-resolution image series shown in Fig. 5d-f was acquired using a Zeiss LSM 980 microscope at 2× digital zoom, 0.05 µm/pixel. Images were captured in super-resolution (SR) mode with a frame acquisition time of 10.52 seconds without any averaging.

Reflected Light Confocal Image Pre-processing

Reflected light confocal images underwent sequential denoising: initial 2D-wavelet-based denoising[77] for spatial noise reduction, followed by FastDVDnet video denoising for temporal noise[78]. FastDVDnet's multi-frame processing effectively reduced noise while maintaining structural integrity.

Deep Learning-Based Segmentation and Auto-Annotation Pipeline

Our study employed an active learning approach to annotate a large dataset of cochlear images, combining manual annotation with an auto-annotation pipeline inspired by Pettersen et. al, 2022[53]. We first annotated approximately 10,000 images derived from 10 distinct time series using the open-source software Microscopy Image Browser[53]. These manually annotated images served as the foundation for training our deep learning model and initiating the auto-annotation process.

We initially annotated nine distinct classes: TM, tectorial membrane; BC, border cells; OHC, outer hair cells; HeC, Hensen's cells; RL, reticular lamina; PC, outer pillar cell head; IHC, inner hair cells/inner pillar cell complex and RM, Reissner’s membrane. However, the RM was excluded from the final analysis due to its dynamic behavior during loud sound stimulation, as it transiently moved in and out of the acquisition frame[47]. For the two-class dataset a single OHC was delineated from the RL and other neighboring structures.

Auto-Annotation Pipeline

The auto-annotation pipeline was developed to efficiently process and annotate large volumes of cochlear images with eight segmentation masks per image. The process began with a training phase where three parallel training scripts were executed, each utilizing different augmentation values. We employed a DeepLabV3Plus architecture with a ResNet18 encoder as the segmentation model. Data augmentation was performed using the Albumentations library and included resizing, horizontal and vertical flipping, rotation, brightness and contrast adjustments, gamma correction, blurring, and grid

distortion. The training process was optimized using mixed precision computations and gradient accumulation to maximize efficiency.

Following the training phase, which consisted of 152 epochs with 133 199 training images and 24 658 validation images for the 7 structure segmentation (Fig. 2, Supplementary Movie 2-5), and 65 epochs with 25 302 training images and 2 016 validation images for the single OHC segmentation (Fig. 5), our pipeline entered an inference stage where the trained model was applied to new, unannotated images. Three inference scripts were executed concurrently, one for the current state of each trained network. Post-inference, a combination script merged the results from three inference runs. This script applied morphological operations such as expansion and erosion to refine the segmentation boundaries. We then employed an adaptive thresholding mechanism to identify high-quality annotations. The intersection over union (IoU) threshold was dynamically adjusted based on the distribution of scores in each iteration, typically ranging between 0.97 and 0.985.

Suppl. Fig. 4 displays the confusion matrix, which quantifies the performance of our classification model by comparing the predicted labels against the true labels for each class.

Quality Control and Filtering

Quality control of auto-annotated images was then applied. Each potential auto-annotation underwent a series of checks to ensure all required anatomical classes were present with a reasonable position and size. We applied various exclusion criteria, including checks for large holes in the segmentation, verification of area ranges for specific anatomical structures, and confirmation of proper spatial relationships between different segmented elements (e.g. the upper border of the OHC class should not be positioned above the RL). Excluded auto-annotated images were moved from the trainset back into the un-annotated folder.

To further refine our dataset, we calculated and analyzed several metrics for each annotation, including centroid Y- and X-positions, lengths, widths and areas. These metrics were visualized through automatically generated graphs, and a three-frame rolling window analysis of the graphs allowed the identification of potential outliers (> 0.3x the rolling average) or inconsistencies in the annotations. Images flagged as outliers were returned to the original image pool for re-annotation.

Dataset Expansion and Model Refinement

Images that successfully passed all quality control measures were incorporated into the training dataset for subsequent iterations. This process allowed for the continuous expansion of our annotated dataset while maintaining high-quality standards. Each iteration potentially added more auto-annotated images to the training set, which in turn were used to further refine the model.

To aid in visual inspection and quality assurance, our pipeline generated overlay images that combined the original cochlear image with the segmentation mask. These overlays provided a quick means of assessing annotation accuracy.

The entire auto-annotation process was designed to run iteratively, with each cycle potentially improving annotation quality and expanding the dataset. Extensive error handling and logging mechanisms were integrated throughout the pipeline to track progress and swiftly identify any issues that might arise.

Evaluation and Validation

To evaluate the accuracy and reliability of our semantic segmentation network, we implemented a rigorous evaluation procedure. Approximately 5 % of fully annotated time series were set aside as a validation set, excluded entirely from the training process during the active learning auto-annotation process. Further, the final fully annotated dataset was split with an approximate 85%/15% training/validation split (split randomly by full timeseries, such that no frames from the same timeseries were present in both the training and validation set) and trained from scratch. This enabled independent evaluation of the model's performance in terms of structure-specific and overall Intersection over Union (IoU) and accuracy metrics.

Our evaluation results revealed high performance metrics for the CNN-based structure detection. We achieved a mean IoU of 0.913, a mean pixel accuracy of 0.987, demonstrating the semantic segmentation model's ability to generalize robustly and accurately detect and delineating the cochlear structures on unseen timeseries.

Metric quantification

Metric analysis involved computing several metrics including X and Y centroid coordinates, lengths, widths and areas for each segmented structure. For 412-image series (Pipeline Suppl. Fig. 1), baseline correction used the median of the initial 63 data points at 0 dB SPL as explained in the Results section.

For time-resolved series consisting of 12 images (captured during a 250 Hz, 70 dB SPL tone burst, Fig. 5), the data processing involved two sequential steps. First, each measurement series was baseline-corrected by subtracting its median value. Then, temporal smoothing was applied using a 3-point moving average window, where each data point was averaged with its two adjacent points, helping to reduce noise while preserving the underlying temporal dynamics. The complete time series spanned 4 ms with 12 equidistant measurement points.

Statistical Analysis

Linear Mixed Model (LMM) analysis was performed to investigate time-dependent responses across experimental conditions, using MATLAB’s fitlme function. The model incorporated fixed effects for condition and time, with a random intercept for subject to account for repeated measures and inter-subject variability. Data from specific time windows were analyzed, with 140–180 ms used for fast adaptation and 126–183 ms for

slow electrical and mechanical adaptation. This ensured that the analysis captured the relevant dynamics for each adaptation process.

Data from individual subjects were concatenated, and a design matrix was constructed to include condition, subject identity, and time as variables. The dependent variable (response) and independent variables were transformed into a table format suitable for LMM analysis, with condition and subject identity treated as categorical variables. The model used Restricted Maximum Likelihood (REML), the default and preferred method for mixed-effects models, ensuring accurate and robust parameter estimation.

The model, specified as $\text{Response} \sim \text{Condition} + \text{Time} + (1 \mid \text{SubjectID})$, enabled evaluation of condition and time effects while accounting for variability between subjects. Pairwise comparisons were conducted between selected conditions, including control vs. salicylate, control vs. EGTA, and control vs. DHS. Key outputs included fixed effects coefficients, providing estimates for condition and time effects, and ANOVA results, offering statistical significance for the model terms. To ensure the reliability of the results, the analysis accounted for hierarchical data structures and provided detailed insights into the response trajectories under different experimental conditions. This approach ensured robust statistical comparisons while capturing the dynamics of both fast and slow adaptation processes.

All references cited in the Supplementary Materials and Methods are included in the main manuscript's reference list.