Brain Stimulation 16 (2023) 1486–1500

Contents lists available at ScienceDirect

Brain Stimulation

journal homepage: www.journals.elsevier.com/brain-stimulation

Devising a framework of optogenetic coding in the auditory pathway: Insights from auditory midbrain recordings

Maria Michael [a,1], Bettina Julia Wolf [a,b,c], Astrid Klinge-Strahl [a,d], Marcus Jeschke [a,b,e], Tobias Moser [a,b,c,f,**], Alexander Dieter [a,g,h,*]

[a] Institute for Auditory Neuroscience and InnerEarLab, University Medical Center Göttingen, 37075, Göttingen, Germany
[b] Auditory Neuroscience and Optogenetics Laboratory, German Primate Center, 37077, Göttingen, Germany
[c] Cluster of Excellence "Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells" (MBExC), University of Göttingen, 37075, Göttingen, Germany
[d] Department of Otolaryngology, University Medical Center Göttingen, 37075, Göttingen, Germany
[e] Cognitive Hearing in Primates (CHiP) Group, German Primate Center, 37077, Göttingen, Germany
[f] Auditory Neuroscience and Synaptic Nanophysiology Group, Max Planck Institute for Multidisciplinary Science, Göttingen, Germany
[g] Göttingen Graduate Center for Neurosciences, Biophysic, and Molecular Biosciences, 37077, Göttingen, Germany
[h] Department of Neurophysiology, MCTN, Medical Faculty Mannheim, Heidelberg University, 68167, Mannheim, Germany

ARTICLE INFO

Keywords:
Hearing restoration
Cochlear implant
Gene therapy
Optogenetics
Sound encoding
Inferior colliculus

ABSTRACT

Cochlear implants (CIs) restore activity in the deafened auditory system via electrical stimulation of the auditory nerve. As the spread of electric current in biological tissues is rather broad, the spectral information provided by electrical CIs is limited. Optogenetic stimulation of the auditory nerve has been suggested for artificial sound coding with improved spectral selectivity, as light can be conveniently confined in space. Yet, the foundations for optogenetic sound coding strategies remain to be established. Here, we parametrized stimulus-response-relationships of the auditory pathway in gerbils for optogenetic stimulation. Upon activation of the auditory pathway by waveguide-based optogenetic stimulation of the spiral ganglion, we recorded neuronal activity of the auditory midbrain, in which neural representations of spectral, temporal, and intensity information can be found. Screening a wide range of optical stimuli and taking the properties of optical CI emitters into account, we aimed to optimize stimulus paradigms for potent and energy-efficient activation of the auditory pathway. We report that efficient optogenetic coding builds on neural integration of millisecond stimuli built from microsecond light pulses, which optimally accommodate power-efficient laser diode operation. Moreover, we performed an activity-level-dependent comparison of optogenetic and acoustic stimulation in order to estimate the dynamic range and the maximal stimulation intensity amenable to single channel optogenetic sound encoding, and indicate that it complies well with speech comprehension in a typical conversation (65 dB). Our results provide a first framework for the development of coding strategies for future optogenetic hearing restoration.

1. Introduction

Electrical cochlear implants (eCIs) are the standard of care to restore hearing in patients suffering from profound sensorineural hearing loss or deafness [1,2]. Since the first implantation in 1957, approximately a million patients worldwide received eCIs, enabling open speech comprehension in most of them [2–4]. Unfortunately, the conductive nature of intra-cochlear fluids causes relatively wide spread of the stimulation current from each electrode contact [5]. In consequence, the spatial selectivity of artificial stimulation of spiral ganglion neurons (SGNs; i.e. the neurons composing the auditory nerve) is limited, and the frequency resolution of electrical hearing restoration is restricted to approximately ten perceptually different stimulation channels [6–8]. As a result, there remains a major unmet medical need for improved hearing restoration, in particular for better understanding of speech in background noise [9,10].

In recent years, optogenetic stimulation of SGNs via optical cochlear implants (oCIs) has been suggested as an alternative mode of hearing

* Corresponding author. Institute for Auditory Neuroscience and InnerEarLab, University Medical Center Göttingen, 37075, Göttingen, Germany.
** Corresponding author. Institute for Auditory Neuroscience and InnerEarLab, University Medical Center Göttingen, 37075, Göttingen, Germany.
E-mail addresses: tmoser@gwdg.de (T. Moser), alexander.dieter@medma.uni-heidelberg.de (A. Dieter).
[1] Present address: Department of Oto-Rhino-Laryngology, University Medical Center Freiburg, Kilianstraße 5, 79106 Freiburg, Germany.

https://doi.org/10.1016/j.brs.2023.09.018
Received 12 June 2023; Received in revised form 19 September 2023; Accepted 21 September 2023
Available online 29 September 2023

restoration. Light can be better confined in space and hence activates SGNs with greater spectral selectivity as compared to electrical current [3,11–13]. Studies in rodents employing virus-mediated expression of tailored channelrhodopsins and electrophysiological recordings have proven the feasibility of optogenetic SGN-stimulation with higher spectral selectivity [12,14–16], broader dynamic range [17,18], and promising temporal fidelity [17–21]. Furthermore, the perception of these stimuli has been demonstrated in gerbils and rats on the behavioral level [16,22], and restoration of auditory function by cochlear optogenetics has been demonstrated in deafened mice, gerbils, and rats [12,16,20,22]. In first attempts towards clinical translation, rigorous functional and histological analyses of various viral constructs for optogenetic transduction have been performed [17,19,23,24], and morphological characterization of the native and implanted cochlea in several species has been done to guide the development of future oCIs [25], while modelling studies predicted improved hearing restoration by optical as compared to electrical CIs [26].

Finally, microsystems engineering of oCIs based on micrometer-scaled light-emitting diodes (μLEDs) [27,28] has enabled multi-channel optogenetic SGN stimulation in anesthetized gerbils with good spectral selectivity [15]. Furthermore, another implementation of LED-based multichannel oCIs [29] has been used to stimulate the SGNs of rats and gerbils, confirming increased spectral resolution as compared to eCIs as well as behaviorally relevant perception of optogenetic stimulation of the auditory pathway [16]. Recently, the preclinical development of optogenetic hearing restoration has been complemented by the engineering of a complete CI-system, which is capable of processing acoustic stimuli in close-to-real-time and driving both electrical and optical multi-channel cochlear implants in small animals [30].

Yet, devising strategies for optogenetic sound encoding remains as a key step towards preparing the clinical translation of optogenetic hearing restoration. This task requires careful parametrization of optogenetic SGN stimulation regarding coding of spectral, temporal, and intensity information. Properties of individual pulses should be estimated, while the assembly of pulses into stimulation patterns that mimic natural sound encoding should consider an increase of stimulation channels as well as an opsin-dependent temporal structure [13,30]. In addition, the coding strategy should ideally accommodate the properties of the semiconductor emitters and their driver architecture. Specifically, laser diodes accommodate nanosecond timing and operate most efficiently when driven with brief and high amplitude currents. Another motivation for employing ultrashort pulses could arise from the logic of addressing emitter arrays, which might involve multiplexing of current sources and matrix addressing [27]. Finally, the coding strategy should aim to minimize the power budget of optical CIs.

In this study, we systematically explored the parameter space for optogenetic stimulation in terms of intensity and temporal structure in order to define suitable quantal stimuli for optical sound encoding. We then varied the stimulus' duty cycle, and identified stimulation patterns that would both optimize energy efficiency of optical sound encoding without significantly affecting the reliability of neural activation, and flexibly accommodate the operation of optoelectronic stimulation hardware [30]. Furthermore, we estimated the ability of optogenetically modified SGNs to encode temporal gaps in stimulation patterns. Gap detection results, on the one hand, can be used to estimate the temporal acuity of optical sound encoding, while on the other hand they serve as an indicator of temporal resolution in speech processing [31,32]. Finally, we compared activity that was optically evoked by a single waveguide with corresponding activity evoked by acoustic stimulation in order to approximate the dynamic range and the maximum stimulus intensity that can be expected from optogenetic hearing restoration by a single emitter. Taken together, the results of this study will guide the development of coding strategies for optical hearing restoration, balancing biomimetic stimulation parameters and energy consumption of future oCIs.

2. Results

In this study, we investigated the stimulus-response characteristics of the auditory pathway upon single fiber-based optogenetic stimulation of the auditory nerve in Mongolian gerbils with a wide range of optical pulses differing in intensity, duration, and repetition rate (Fig. 1A). To this end, we injected a suspension of adeno-associated virus (AAV-PHP.B [33] [4.6e12 gc/ml] or AAV-PHP.eB [34] [3.8–4.3e12 gc/ml]) carrying DNA encoding for the calcium translocating channelrhodopsin *CatCh* (a variant of the blue-light activated ChR2 [35]) fused to enhanced yellow fluorescent protein (eYFP) under the control of the neural human synapsin promoter into the modiolus of adult gerbils [22], leading to *CatCh*-eYFP expression in SGNs (Fig. 1B-C). Successful optogenetic activation of SGNs was evaluated earliest six weeks after injection by recording auditory brainstem responses evoked by blue light pulses delivered to the cochlea (oABR) [22] via a laser-coupled optical fiber placed in the round window and oriented to illuminate large parts of the cochlea (Fig. 1D, Fig. S1). This specific fiber orientation was chosen to maximize the cochlear spread of excitation which enables broad activation of the auditory midbrain across different frequency laminae, when high stimulation intensities are chosen [14,15]. In the AAV-injected cochleae of animals which showed a visually detectable and reproducible oABR, we found 21.6% of the SGNs averaged across all tonotopic places to express *CatCh*-eYFP (Fig. 1C, $n = 7$). No significant difference in SGN density was observed between injected and non-injected cochleae (2.43 ± 0.8 vs. 2.9 ± 0.52 SGNs/1000 μm^2, $p = 0.16$, two-sample *t*-test; $n = 24/8$ cochlear turns from $N = 7/4$ animals), indicating no significant adverse effects due to the injection, AAV-transduction or *CatCh*-eYFP expression (Fig. S2). In a subset of analyzed animals (5/7), sporadic opsin expression was also observed in inner hair cells, although they do not natively express synapsin [36] (Fig. S3).

We then inserted a linear 32-channel silicon electrode array along the central nucleus of the inferior colliculus (ICC) of the auditory midbrain in order to record spiking activity of neuronal ensembles (multi-unit-activity; Fig. 1E). Electrode placement was confirmed by acoustic stimulation, which reveals the tonotopic organization of the ICC (4.23 ± 0.79 octaves/mm (mean ± standard deviation); *Pearson's r* $= 0.84$, $p = 6.6 \times 10^{-143}$, Fig. 1F; in line with previous reports e.g. Refs. [14,15,37]). Subsequently, we stimulated the auditory nerve with various optical stimuli, covering a wide range of intensities and temporal structures (30 repetitions were used for each stimulus condition; for a full list of animals contributing to each dataset see Supplementary Table 1). Subsequent analysis of auditory midbrain activity – which represents spectral, temporal and intensity information – evoked by physiological or artificial (bionic) cochlear stimulation then allowed us to draw conclusions about different aspects of optogenetic auditory nerve activation and potential implications for future strategies of optical sound encoding.

Properties of auditory midbrain responses evoked by individual light pulses in the cochlea: First, we presented individual light pulses with durations ranging from 10 μs to 10 ms and intensities (radiant flux) ranging from 2 to 32 mW at low repetition rates (2 Hz), in order to evaluate the response characteristics for cochlear stimulation with individual light pulses (Fig. 2A and B; see Fig. S4 for pulse shapes). The response strength (given as the discrimination index d', calculated based on the spike rates) of midbrain neurons increased with pulse intensity at a given pulse duration (*Pearson's r* $= 0.38/0.47/0.44/0.36/0.39/0.34$, $p < 1 \times 10^{-26}$ for pulse durations of 0.2/0.5/1/2/4/10 ms respectively; $n/N = 192$ units/6 animals; Fig. 2C). Moreover, response strength increased with pulse duration at a given pulse intensity ($r = 0.46/0.51/0.52/0.60/0.56$, $p < 1 \times 10^{-30}$, for intensities of 2/4/8/16/32 mW, respectively; $n/N = 192$ units/6 animals Fig. 2D). Responses were specific to optogenetic stimulation, as they were not observed in non-injected control animals using even longer pulse durations and higher intensities (Fig. S5). Both pulse duration and pulse intensity affected the

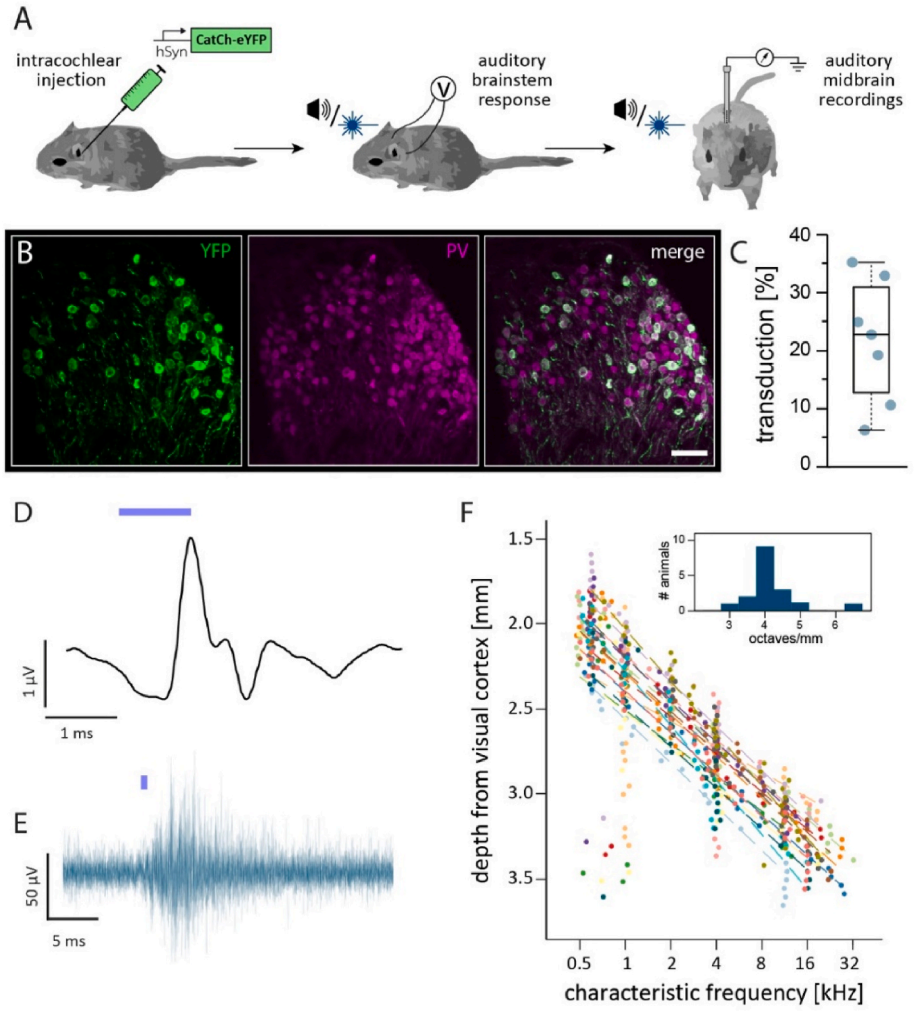

Fig. 1. Optogenetic activation of the auditory pathway. (A) Scheme of experimental workflow. **(B)** *CatCh*-eYFP transduced (green) SGNs (identified by parvalbumin-expression, magenta) in the middle turn of an exemplary AAV-injected cochlea. Scale bar: 50 µm. **(C)** Fraction of *CatCh*-eYFP expressing SGNs in AAV-transduced cochleae of oABR-positive animals. **(D)** Optogenetically evoked auditory brainstem response (mean of 1000 repetitions) in an exemplary *CatCh*-expressing gerbil, evoked by a 1 ms light pulse (indicated by the blue bar) of ~40 mW, delivered by a laser-coupled fiber placed at the round window. **(E)** 30 voltage traces of a multi-unit recording in the central nucleus of the auditory midbrain in response to a 1 ms pulse of ~32 mW. **(F)** Tonotopic slopes, derived from characteristic frequencies and recording depth ($n = 534$ multi-units, $N = 17$ animals; *Pearson's* $r = 0.84$, $p = 6.6 \times 10^{-143}$). Different colors indicate different animals. Data points in the lower left corner of the graph likely represent neuronal clusters ventral to the central nucleus of the IC [37]. Insert: Distribution of tonotopic slopes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

strength of neuronal responses in injected animals: responses were strongest for stimuli of long duration and high intensity (i.e. high radiant energy; Fig. 2E).

Next, we estimated the cochlear spread of excitation by the number of responsive electrodes as a function of pulse duration and pulse intensity. Similar to the response strength, also the recruitment of responding multi-neuronal clusters (i.e. responding electrodes, defined as electrodes which reached a d' ≥ 1) increased with pulse intensity ($r =$ 0.61/0.59/0.61/0.54/0.54/0.54, $p < 0.01$, for 0.2/0.5/1/2/4/10 ms pulse duration respectively; $N = 6$ animals; Fig. 2F) and duration ($r =$ 0.59/0.61/0.58/0.67/0.65, $p < 1 \times 10^{-20}$, for intensities of 2/4/8/16/32 mW, respectively; $N = 6$ animals; Fig. 2G). Again, both pulse duration and pulse intensity affected the number of responsive electrodes, which scaled with the radiant energy of a given pulse (Fig. 2H). To reach a certain response strength and/or spatial extent of midbrain activity, pulse duration could be compensated by pulse intensity and vice versa, resulting in similar levels of response strength and spread of excitation, at least within a certain range, as indicated by the iso-contour-lines (Fig. 2E, H).

Optogenetic stimulation with pulses of equal radiant energy: To further investigate the interaction between durations and intensities of individual pulses, we again presented individual pulses, but this time durations and intensities were balanced to result in stimuli of equal radiant energy (2, 4, 6, 8, and 10 µJ). As shown above, the response strength increased as a function of radiant energy, suggesting synergistic effects of pulse duration and pulse intensity (Fig. 3A, $N = 5$ animals). Comparing the response strength across stimuli with identical radiant energy revealed that responses were stronger for short, high intensity pulses as compared to longer pulses of lower intensity (Fig. 3B and C). Across all energy levels tested, the maximum response strength for pulses with the highest intensity and shortest duration was significantly larger than the maximum response strength for pulses with the lowest intensity but longest duration (Fig. 3D; $p < 0.001$; Wilcoxon rank sum test with Bonferroni correction; only units that were responsive to at least one intensity-duration-combination were considered, $N = 5$ animals). Thus, for energy efficient coding, short pulses of high intensity should be preferred over longer pulses with lower intensities, which suits most efficient operation of laser diodes.

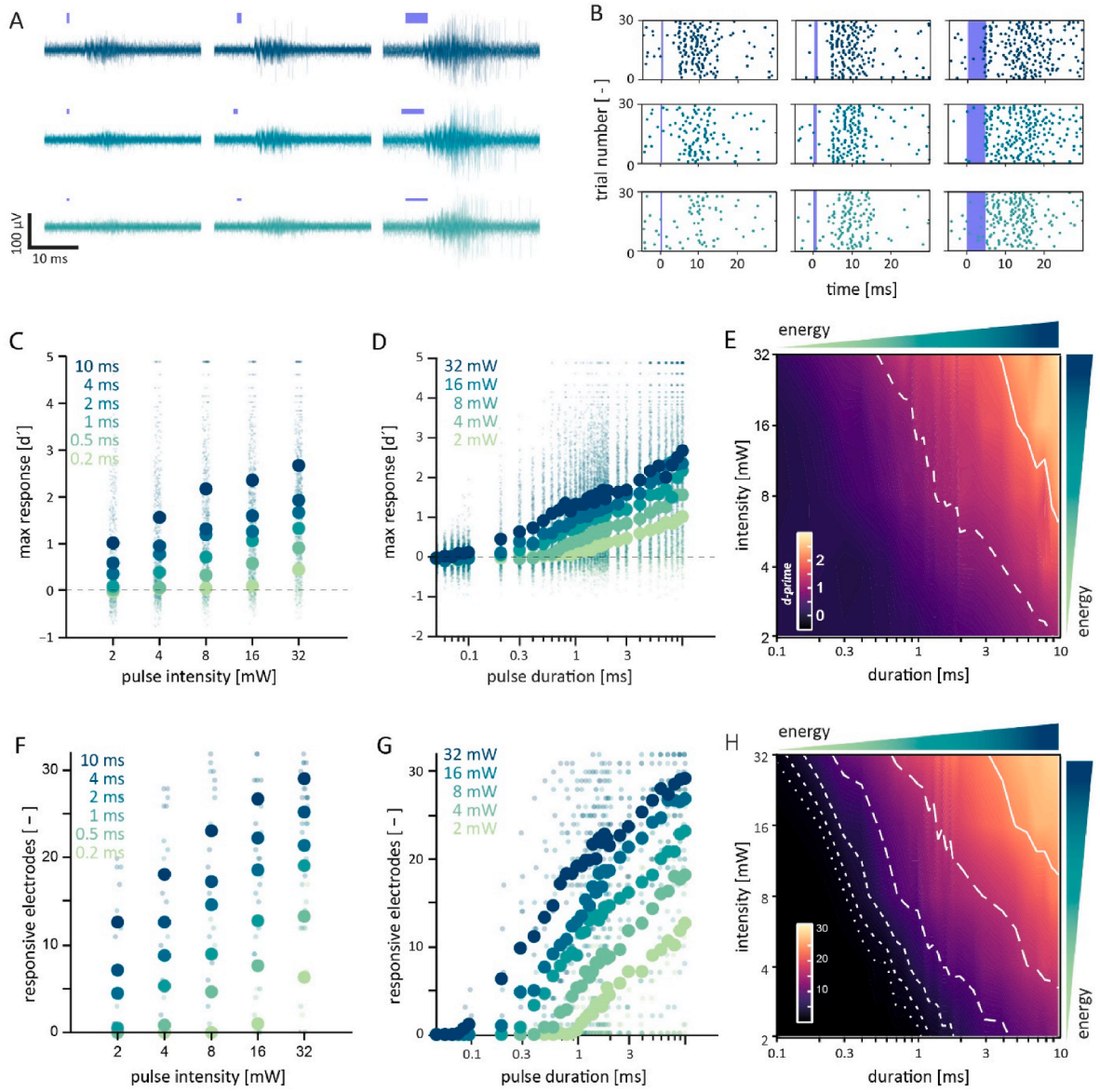

Fig. 2. Dependence of auditory midbrain responses on pulse intensity, pulse duration, and radiant energy. (A) 30 individual voltage traces of an exemplary multi-unit recording in the auditory midbrain in response to optogenetic auditory nerve stimulation. Stimuli – indicated by the blue bars – had durations of 0.5, 1, and 5 ms (left to right) and intensities of 32, 16, and 8 mW (top to bottom). Pulse duration is indicated by bar width, pulse intensity by bar height. **(B)** Scatter plots derived from the traces in (A). **(C)** Response strength as a function of stimulus intensity, shown for pulses of 0.2, 0.5, 1, 2, 4, and 10 ms duration ($n = 192$ units from $N = 6$ animals). Large symbols indicate mean. **(D)** Response strength as a function of stimulus duration, shown for stimuli of 2, 4, 8, 16, and 32 mW intensity ($n = 192$ units from $N = 6$ animals). **(E)** Heat map indicating the response strengths as a function of both pulse duration and pulse intensity ($N = 6$ animals). Iso-contour lines are indicated for a d' of 1 and 2. **(F)** Number of responsive electrodes as a function of stimulus intensity, shown for pulses of 0.2, 0.5, 1, 2, 4, and 10 ms duration ($N = 6$ animals). Large symbols indicate mean. **(G)** Responding units as a function of stimulus duration, shown for stimuli of 2, 4, 8, 16, and 32 mW intensity ($N = 6$ animals). **(H)** Heat map indicating the number of responding electrodes as a function of both pulse duration and pulse intensity ($N = 6$ animals). Iso-contour lines are indicated for 1, 2, 4, 8, 16, and 24 active electrodes, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

As shorter pulses drove SGNs more efficiently, we next determined the absolute pulse duration thresholds of individual pulses which could still elicit significant neuronal responses (d' $\geq$ 1) at various laser intensities. Pulses of higher intensity recruited more multi-neuronal clusters when pulse duration was constant (Fig. S6A) and had lower pulse duration thresholds (Fig. S6B) as compared to pulses with low intensity. As pulses of higher intensity recruited more units, we focused our analysis on units which were responsive even to the lowest laser intensity presented (2 mW) and found that their median pulse duration threshold decreased with increasing stimulation intensity (2.5/1.1/0.6/0.5/0.2 ms pulse duration for intensities of 2/4/8/16/32 mW, respectively; $n = 68$ units from $N = 8$ animals; Fig. S6B). The shortest pulse duration eliciting a significant response in any multi-unit was 80 µs (at 32 mW).

Stimulation of the auditory nerve using sub-threshold pulselets: Next, we tried to evoke neuronal responses with ultrashort, sub-threshold pulses in the microsecond range. These stimuli, which we have termed pulselets, were defined as pulses which were not capable of driving SGNs when presented individually. Hence, we assembled individual pulses from these pulselets by repeating them at stimulation rates in the kHz-range (i.e. well beyond the physiological limits of both SGNs [38] and *CatCh* [35]). Specifically, we created pulses with a total duration of 1 ms

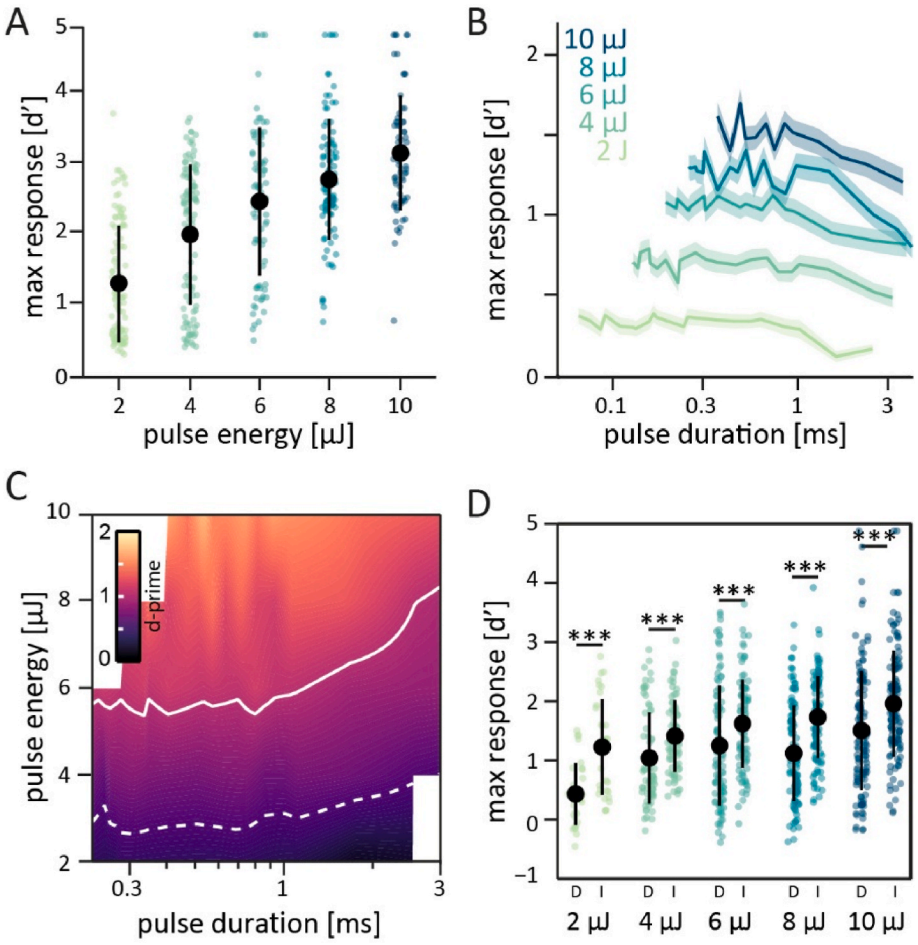

Fig. 3. Dependence of auditory midbrain responses on duration-intensity-combinations at equal radiant energy. (A) Maximum response strength for duration-intensity combinations equivalent to a given pulse energy (mean ± SD; maximum response of $n = 160$ units from $N = 5$ animals). **(B)** Maximum response strength as a function of pulse duration. Intensity is counterbalanced to reach the respective radiant energy of 2, 4, 6, 8, or 10 μJ (color-coded). Data depicts mean ± SEM of all multi-units recorded ($n = 160$ units from $N = 5$ animals**). (C)** Color coded maximum response strength as a function of pulse duration (and – in an inverse relationship – pulse intensity), at different pulse energies. Iso-contour lines are indicated for a d' of 0.5 and 1. Data depicts mean of all multi-units ($n = 160$ units from $N = 5$ animals). **(D)** Maximum response strength of multi-units elicited by pulses with the longest duration but lowest intensity (*D*) versus pulses with the shortest duration but highest intensity (*I*) at different radiant energy. Data depicts mean ± SD of responsive (i.e. reaching a d' > 1 in response to at least one duration-intensity-combination) units (n = 32/71/99/113/127 units for 2/4/6/8/10 μJ, respectively; recorded from $N = 5$ animals). *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$ (Bonferroni-corrected, signed Wilcoxon rank sum test). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

assembled from pulselets of 20 μs with repetition rates of 25 or 12.5 kHz, resulting in a duty cycle of 50 or 25% (Fig. 4A). We then compared the responses of these stimuli to the ones upon a single pulse of 1 ms duration (i.e. a duty cycle of 100%).

Presented at these high stimulation rates, pulselets of 20 μs could indeed evoke neuronal responses when assembled into a 1 ms pulse with a duty cycle of 50 and 25% (Fig. 4B and C). Comparing the maximum response strength of multi-units recorded from responsive electrodes (i.e. electrodes reaching a d' > 1 in response to at least 3 different intensities of continuous pulses), we observed that the maximum response strength scaled with the duty cycle, i.e. pulses of higher energy evoked stronger responses (d' = 3.2 ± 1.1/2.8 ± 1.2/2.2 ± 1.2 for duty cycles of 100/50/25%, $p = 1.3 \times 10^{-12}$, $F_{483} = 28.97$, one-way-ANOVA; n = 162 units of $N = 6$ animals, Fig. 4D). However, when matching stimulus intensities to result in identical radiant energy of pulses, we observed that stimuli assembled from pulselets evoked stronger responses as compared to continuous pulses (d' = 2.3 ± 0.6/2.7 ± 0.9/3.1 ± 1.0 for duty cycles of 100/50/25%, $p = 4.2 \times 10^{-6}$, $F_{207} = 13.14$, one-way-ANOVA; $n = 70$ units of $N = 6$ animals, Fig. 4E). Furthermore, the energy threshold of multi-unit activity, i.e. the radiant energy required to elicit a significant response (d' ≥ 1), was lower for stimuli assembled from pulselets as compared to a single pulse (7.6 ± 4.6/5.8 ± 2.9/5.1 ± 2.4 μJ for duty cycles of 100/50/25%, $p = 1.9 \times 10^{-7}$, $F_{342} = 16.16$, one-way-ANOVA; n = 115 units of $N = 6$ animals, Fig. 4F and G). These results demonstrate the possibility to grade stimulus intensity by scaling the duty cycle within the stimuli that are constructed from sub-threshold pulselets at kHz rates well exceeding physiological SGN firing rates.

Temporal acuity of optogenetically evoked auditory midbrain responses: High (sub-millisecond) temporal precision is one hallmark of sound encoding. This could be at risk in optogenetic coding, given that channelrhodopsins have off-kinetics in the milliseconds range. How this affects the processing of temporal stimulus properties beyond the auditory nerve has not been reported before. Hence, we aimed to investigate the encoding of temporal properties in the IC upon optogenetic SGN stimulation. To evaluate the temporal acuity of optogenetic sound encoding, we used the gap detection paradigm, for which a temporal gap of varying length is introduced into an ongoing stimulus (Fig. 5A). This measure is of great interest, as gap detection thresholds are a good proxy for temporal resolution in speech perception [31,32]. Furthermore, the ability to detect gaps in an ongoing stimulus would inform us whether pulses assembled from high-frequency pulselets (as shown above) would still be interpreted as a single stimulus, or whether this temporal

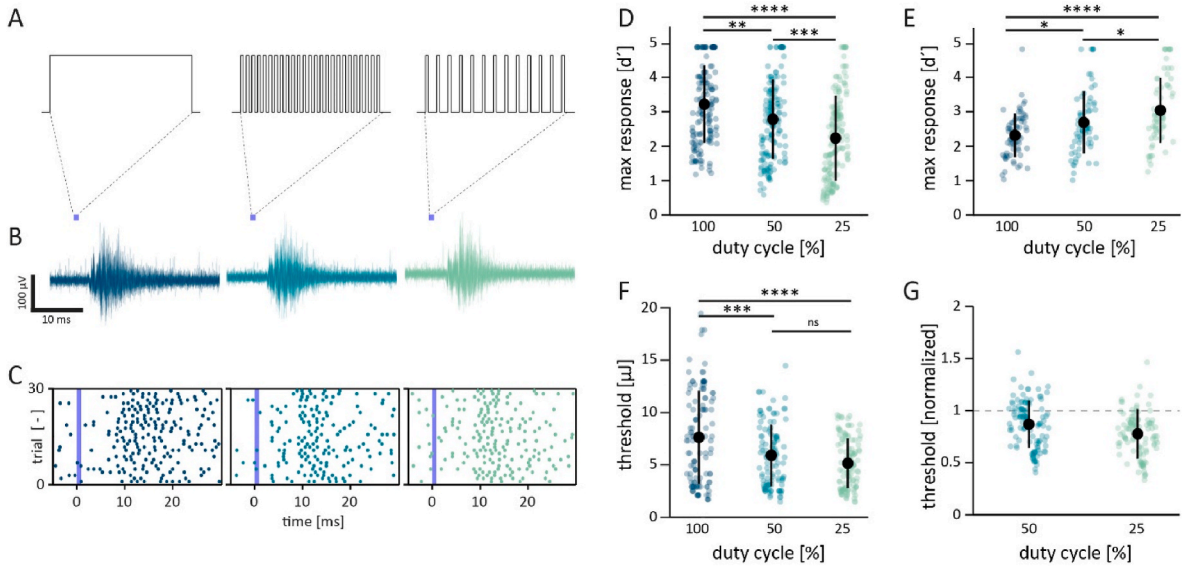

Fig. 4. High-frequency stimulation of the auditory nerve with microsecond **pulselets. (A)** Illustration of 1 ms pulses at a duty cycle of 100, 50, and 25% (from left to right), assembled from 20 µs pulselets at 25 kHz (mid, 50% duty cycle) and 12.5 kHz (right, 25% duty cycle). **(B)** 30 individual traces of an IC multi-unit in response to optogenetic stimulation of the spiral ganglion with different duty cycles. Stimuli are indicated by blue bars, and detailed in (A). **(C)** Scatter plots derived from the traces in (A). **(D)** Maximum response strength in response to the stimuli indicated above ($n = 162$ units of $N = 6$ animals). **(E)** Maximum response strength as in (D), but compensating stimuli of lower duty cycles with higher stimulation intensity in order to result in stimuli of equal energy ($n = 70$ units of $N = 6$ animals). **(F)** Radiant energy needed to reach threshold (d' ≥ 1) with stimuli of 100, 50, and 25% duty cycle, respectively ($n = 115$ units of $N = 6$ animals). **(G)** Energy needed to reach threshold with pulses of 50 and 25% duty cycle, normalized to a duty cycle of 100%. */**/***/**** indicate $p < 0.05/0.01/0.001/0.0001$, respectively. ns = not significant. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

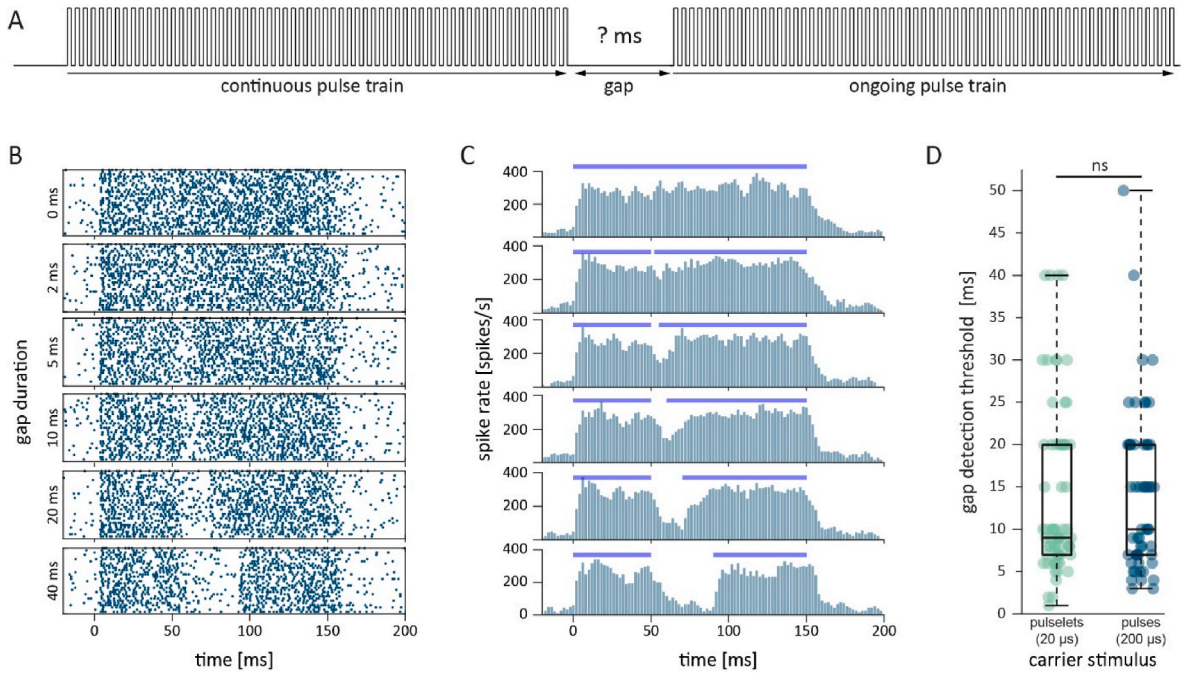

Fig. 5. Gap-detection during ongoing optogenetic stimulation. (A) Scheme of a gap detection stimulus: Gaps of various durations are embedded in an ongoing train of optogenetic pulses or pulselets to probe the temporal acuity of optogenetic auditory nerve stimulation. **(B)** Scatter plot of multi-unit activity ordered from no-gap (top) to gaps of increasing duration. Individual lines depict different trials ($n = 30$ for each stimulus). **(C)** Peri-stimulus time histogram (PSTH) of the multi-units shown in (B). Blue: Pulse train. **(D)** Gap detection thresholds detected in an ongoing stimulus train assembled from 20 µs pulselets repeated at 25 kHz (left) and 200 µs pulses, repeated at 2.5 kHz (right). ns = not significant; Wilcoxon rank sum test, $p = 0.95$, $n = 67/55$ units from $N = 5/5$ animals. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

fine-structure might impact the encoding of single pulses in a more complex way, as in the case of multi-pulse integration of high-frequency electrical stimulation [39].

To evaluate temporal acuity, we used pulse trains (150 ms duration, 30–35 mW), either assembled from sub-threshold pulselets (20 μs duration) or from pulses that also could evoke neuronal responses when presented individually (200 μs duration). Stimulation frequencies were matched to achieve a duty cycle of 50% (i.e. 25 kHz for pulselets of 20 μs, and 2.5 kHz for pulses of 200 μs). Gaps of various durations (including a no-gap condition) were created by omitting pulselets/pulses 50 ms after stimulus onset (Fig. 5B and C). 30 repetitions of each stimulus have been presented in a pseudo-randomized order, and gap detection thresholds (GDT) were then estimated by comparing multi-unit spike rates during the gap window of various gap conditions to the no-gap condition. The first gap duration with a significant decrease in spike rate was then taken as the GDT (given that the stimulus with the next longer gap also showed this decrease).

GDTs obtained from pulse trains assembled from 20/200 μs amounted to 9 ± 3/10 ± 5 ms, respectively (median ± median absolute deviation), while no statistically significant difference was observed between these conditions ($p = 0.95$, Wilcoxon rank sum test; $n = 67/55$ units from $N = 5/5$ animals; Fig. 5D). Hence, for the first time, to our knowledge, we report that temporal gaps of ~9–10 ms can be encoded in an ongoing optogenetic stimulus. This is well above the inter-pulselet interval (20 μs), indicating that pulselets of these stimuli might temporally not be resolved, and hence these stimuli, even though they do not employ continuous illumination, will most likely be perceived as a continuous stimulus.

Temporal response properties of multi-unit activity in the auditory midbrain: Following the estimation of temporal acuity, we turned to investigate the ability of multi-neuronal responses to encode the temporal structure of pulse trains that are sufficient to evoke responses when applied individually, a parameter that likely becomes relevant when considering optogenetic encoding of the temporal structure of real-life acoustic stimuli such as speech. To this end, we presented pulse trains of 101 ms duration, assembled from 1 ms long pulses at 32 mW intensity, with stimulation rates varying between 10 and 500 Hz in steps of 10 Hz (Fig. 6A and B). Readout measures were i) discharge rate, ii) vector strength, and iii) cut-off stimulation rate (i.e. the highest stimulation rate that still showed significant vector strength) of multi-units

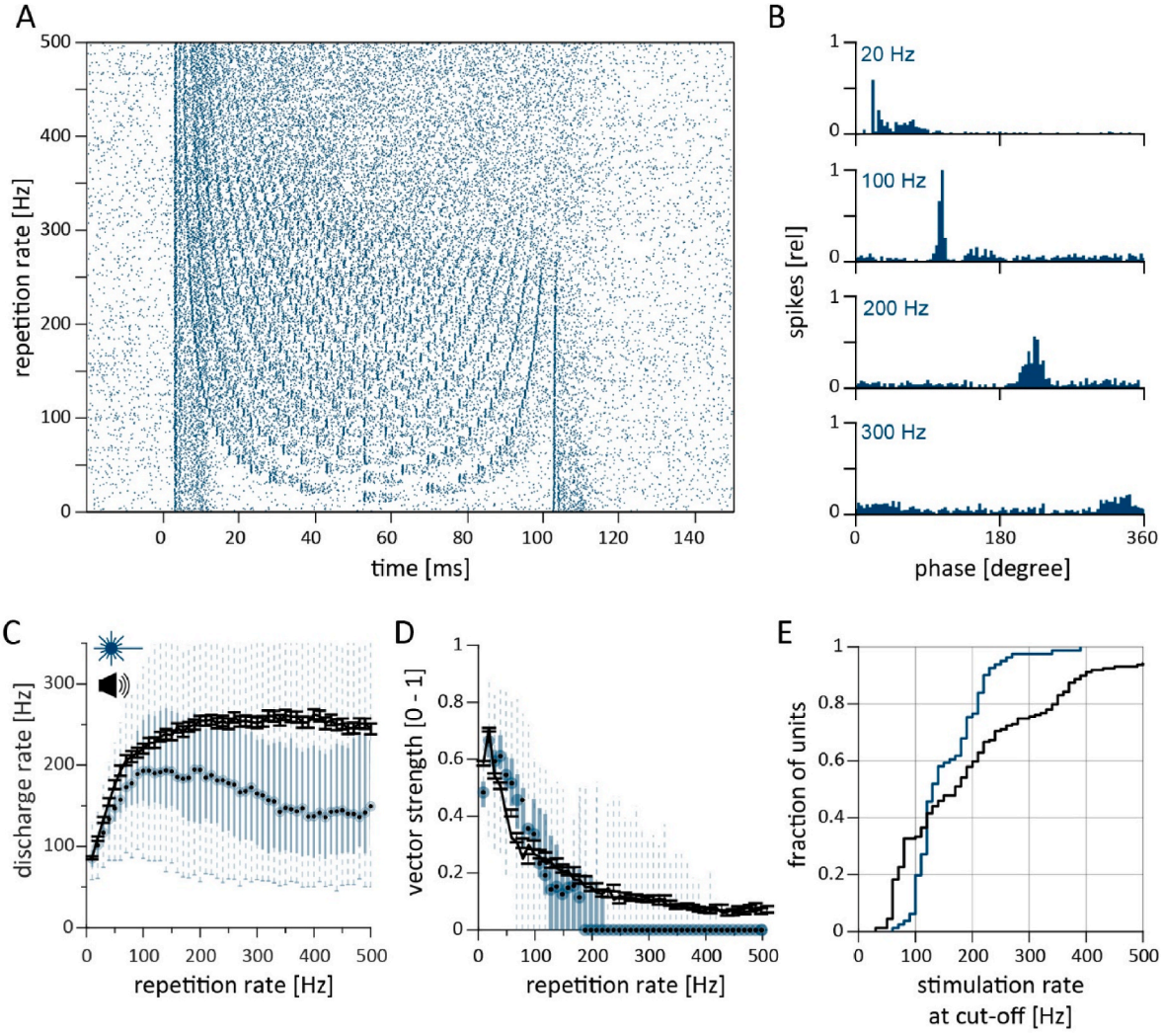

Fig. 6. Temporal properties of multi-unit activity in the inferior colliculus in response to optogenetic SGN stimulation. (A) Exemplary multi-unit in response to 100 ms pulse trains, assembled from pulses of 1 ms duration and 32 mW intensity, with repetition rates varying from 10 to 500 Hz, in 10 Hz steps. **(B)** Phase plots of the unit shown in (A) in response to 20, 100, 200, and 300 Hz, respectively. **(C)** Discharge rate, and **(D)** vector strength in response to 100 ms pulse trains as a function of repetition rate (blue). Boxplots indicate 0-25-50-75-100 percentiles. Black graph shows data of wildtype animals in response to acoustic click trains of ~75 dB at identical repetition rates (mean ± SEM). Data from $n/N = 81/5$ units/animals and $n/N = 159/6$ units/animals for optogenetic and acoustic stimulation, respectively. **(E)** Cumulative density plot assembled from the cut-off frequency for the units shown in (D). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

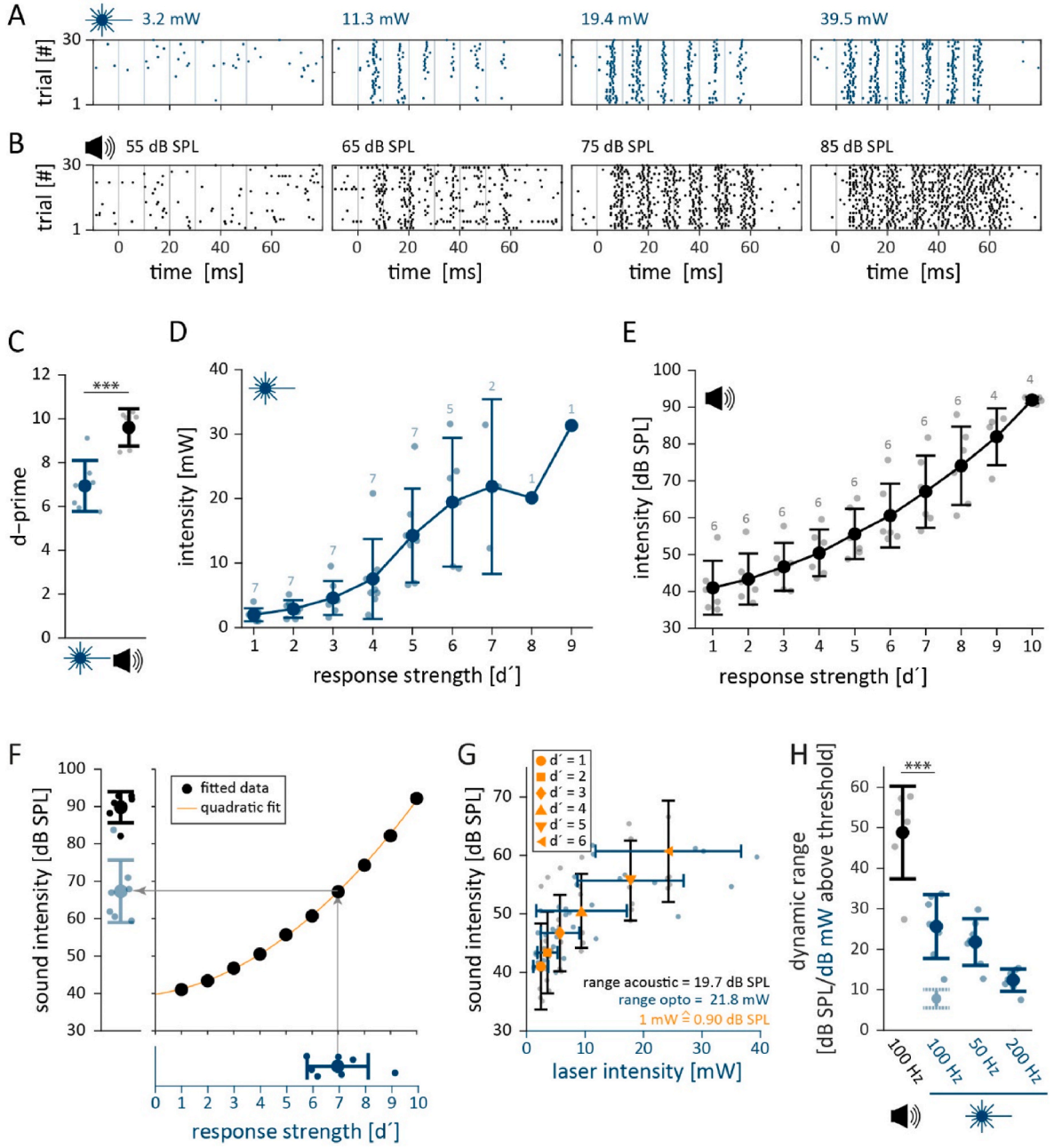

Fig. 7. Estimation of the dynamic range that can be encoded with optogenetic stimulation of the auditory nerve. (A) Scatter plot of an exemplary multi-unit in response to optogenetic pulse trains of increasing laser intensity. **(B)** Scatter plot of an exemplary multi-unit in response to acoustic click trains of increasing sound pressure level. **(C)** Maximum response strength (cumulative d') that could be elicited with 51 ms stimulus trains assembled either from 1 ms laser pulses of varying intensity (left) or 0.3 ms acoustic clicks of varying sound pressure level (right) presented at a stimulation rate of 100 Hz ($n = 7/6$ animals for optogenetic/acoustic stimulation, respectively). **(D)** Laser intensity needed to elicit IC responses of a given strength with 51 ms pulse trains of 1 ms long pulses at 100 Hz in optogenetically transduced animals (mean ± SD: n is indicated for each bar individually). **(E)** Sound pressure level needed to elicit IC responses of a given strength with 51 ms click trains of 0.3 ms long acoustic clicks at 100 Hz in naïve, hearing animals for comparison (mean ± SD; n is indicated for each bar individually). **(F)** A quadratic fit of the sound pressure level needed to reach certain d' values was used to estimate the maximum sound intensity that can be reached with optogenetic stimulation. **(G)** Comparison of stimulus intensities needed to elicit IC responses of a given strength in different sets of animals. Data from (D) and (E) is plotted against each other. **(H)** Estimation of the dynamic range that can be encoded by wildtype animals in response to an acoustic click train (51 ms duration, 100 Hz repetition rate; black) and by optogenetically transduced animals (blue) in response to optical pulse trains (51 ms duration) of 100, 50, and 200 Hz, respectively. Transparent marker: dynamic range (dB mW) as previously estimated [14] (see also Fig. S7). ***: $p < 0.001$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

recorded in the auditory midbrain. The discharge rate of optogenetically driven multi-units increased with stimulation rates up to 100 Hz, where it reached a peak of ~200 spikes/s, before decreasing again to rates of ~150 spikes/s (Fig. 6C). Up to stimulation rates of ~100 Hz, the increase in optogenetically driven discharge rate was similar to the increase in acoustically driven discharge rate recorded in non-injected wildtype animals upon acoustic click trains (i.e. broad band stimuli) of similar temporal structure, presented at 75 dB SPL. However, acoustically driven discharge rates increased further, reaching a plateau of ~250 spikes/s at stimulation rates of approximately 200 Hz (Fig. 6C).

Next, we analyzed the vector strength – a measure of spike synchronization to the stimulus – calculated from spike timing of these units, starting 50 ms after stimulus onset to avoid a bias in favor of the synchronized onset response (Fig. 6B, D). As for acoustically driven firing of midbrain neurons, vector strength of optogenetically driven firing decreased with increasing stimulation rate. At stimulation rates greater than 100 Hz, the decline in vector strength was more pronounced in optogenetically driven units: spike synchronization broke down around ~150 Hz, while some synchronization was found up to 500 Hz in acoustically driven units (Fig. 6D). The lower temporal precision of optogenetically driven neural activity was evident also when comparing the cut-off stimulation rates (i.e. the maximum stimulation rate at which statistically significant phase locking was still observed): The median cut-off rate for optogenetic stimulation amounted to 130 Hz, compared to 170 Hz for acoustic stimulation (Fig. 6E). To verify whether individual neurons in the inferior colliculus are capable of encoding temporal structures above 100 Hz, we have isolated a few single units from our data as a preliminary effort to further scrutinize optogenetically driven activity. Here, we found neurons that robustly encoded temporal structures of up to ~300Hz when driven acoustically and ~150 Hz when driven optogenetically (Supplementary Fig. S7).

Estimating the dynamic range of optogenetic intensity coding: In a final experiment we attempted to estimate the perceived intensity of single channel optogenetic stimulation. To this end, we stimulated the SGNs with pulse trains of 51 ms duration, assembled from pulses of 1 ms duration at a stimulation rate of 100 Hz, at varying stimulation intensities. As a comparison, we recorded data from non-injected gerbils in response to acoustic click trains with similar temporal properties, reasoning that comparable activity levels of the midbrain might predict similar intensity percepts.

Both optogenetic and acoustic stimuli were presented at varying intensities, covering the whole range of potentially relevant stimulus intensities from below-threshold stimulation to high intensities aiming to saturate the neuronal response (Fig. 7A and B; Supplementary Figs. S8A and B). We calculated the dynamic range of stimulus intensities that can be encoded by each multi-unit as the difference in discharge rates between the intensities eliciting threshold (10% above baseline activity) and saturation (90% of the maximum response). The dynamic range in response to acoustic stimulation amounted to 30.1 ± 9.4 dB (peak equivalent [pe] SPL) (mean ± SD; $n = 180$ units, $N = 6$ animals) following $20 \times \log10(I_{90}/I_{10})$, where $I_{90/10}$ are the intensities needed to elicit 90/10% of the maximum response, respectively. For optical stimulation, we have chosen a conservative approach [14], where we transferred the stimulus intensities into dB (mW) using the formula $10 \times \log10(I_{90}/I_{10})$. This yielded a significantly smaller dynamic range of 7.5 ± 2.3 dB (mW) (Fig. S8D; $n = 193$ units, $N = 7$ animals; $p = 3.3 \times 10^{-60}$, Wilcoxon rank sum test). We note that this likely underestimates the dynamic range of optogenetic stimulation, as the mechanism of SGN activation relies on ChR-activation by photon absorption.

Next, we approached the intensity coding amenable to optogenetic stimulation based on the direct comparison of acoustic and optic stimulus intensities which elicit neuronal responses of comparable strength. The reasoning was that similar strength of auditory midbrain activation by acoustic and optogenetic stimulation would likely result in a percept of similar intensity. To this end, we first compared the maximum strength of neuronal responses evoked, which was significantly higher for broad-band acoustic stimulation (d' of 9.6 ± 0.9 vs. 6.9 ± 1.2 mean ± SD for acoustic and optogenetic stimulation, respectively; $p = 0.0007$, two-sample *t*-test; $n = 6/7$ animals for acoustic/optogenetic stimulation; Fig. 7C; absolute spike rates are shown in Figs. S8A–C). We then examined the stimulus intensity required to reach a d' value of a given strength for optogenetic (Fig. 7D) and acoustic stimulation (Fig. 7E), respectively. Subsequently, the upper stimulation intensity that optogenetic SGN stimulation could reach was estimated in the following way: First, we fitted the mean sound pressure level required to reach auditory midbrain activation of a given strength in naïve, wildtype animals, using a second-degree polynomial fit (i.e. quadratic; Fig. 7F, top right panel). We then used this fit to translate the maximum response strength achieved by optogenetic stimulation in *CatCh*-injected animals (Fig. 7F, bottom) into the sound pressure level which led to the same response strength in acoustically stimulated animals (Fig. 7F, grey arrows). Using this approach, we concluded that a response strength corresponding to 67.4 ± 8.4 dB SPL can be reached by optogenetic SGN stimulation from a single emitter (Fig. 7F, left panel, blue). For acoustic stimulation, responses increased up to a level of 89.8 ± 4.2 dB SPL (Fig. 7F, left panel, black). However, it is important to note that this value was ceiled by the experimental design, as our system only allowed stimulation up to 93 dB SPL.

In a different approach we aimed to estimate the dynamic range of optogenetic stimulation in dB (SPL) by matching optical stimulus intensities needed to elicit a given IC response strength achieved by acoustic stimulation (Fig. 7G). For a conservative approach, we limited this to a d' of 6, which was the median maximum cumulative d' value for optogenetic responses, in order to not bias the analysis into the direction of animals showing relatively strong responses. In order to cover response strengths from threshold (d' = 1) up to a d' of 6, optogenetic stimuli of intensities ranging from 2.4 to 24.2 mW were needed, and acoustic stimuli required sound pressure levels ranging from 41.0 to 60.7 dB (pe SPL). Hence, the intensity ranges needed to encode a similar range of response strengths were 21.8 mW and 19.7 dB (pe SPL), respectively, resulting in a "conversion factor" of 0.9 dB SPL corresponding to 1 mW. Next, we aimed to estimate the range of stimulus intensities of each modality that can be encoded by the strength of activation in the auditory midbrain. For acoustic stimulation this is quite straightforward, as it simply depicts the difference between the intensity needed to reach a d' of 1 and the intensity needed to reach the maximum d' value that was recorded in each animal. Here we chose the intensity needed to reach the maximum integer d' value and estimated a dynamic range of 48.8 ± 11.4 dB (pe SPL) (mean ± SD; $n = 6$ animals) that can be encoded in response to acoustic click trains (Fig. 7H, black) within the limits of our stimulation set-up. Analysis for optogenetic stimulation was not as trivial: there, we have taken the difference in stimulus intensities (in mW) needed to reach a d' value of 1 and the highest integer d' value (2.4 ± 1.3 and 30.7 ± 9.1 mW; $n = 7$ animals), spanning a range of 28.4 ± 8.7 mW. In order to estimate a dynamic range comparable to acoustic stimulation (i.e. in dB SPL), we have then converted this range of intensities (in mW) with the conversion factor derived from the intensities needed to elicit a certain response strength before (Fig. 7G), and estimated a dynamic range of optogenetic stimulation equivalent to ~25.6 ± 7.9 dB (pe SPL) (Fig. 7H, blue). Even though this estimate for dynamic range that can be optogenetically encoded is still significantly lower than the dynamic range for acoustic stimulation ($p = 0.0012$, Wilcoxon rank sum test), it by far exceeds the previously reported values, both for optical [14] and for electrical stimulation [14,40,41]. Similar results were obtained for stimulation rates of 50 Hz (21.9 ± 5.8 dB, n = 7; $p = 0.29$, paired *t*-test against optogenetic stimulation at 100 Hz), but the dynamic range decreased when using stimulation rates of 200 Hz (12.4 ± 2.7 dB, $n = 7$; $p = 0.002$, paired *t*-test against optogenetic stimulation at 100 Hz).

3. Discussion

In the current study we parametrized the relationship between the properties of optogenetic stimulation of the auditory nerve and the responses of multi-neuronal clusters in the auditory midbrain, a convenient readout of activity in the ascending auditory pathway. Stimulus properties were chosen to cover a wide range of pulse durations, intensities, and temporal structures. Analysis addressed both responses to individual pulses and pulse trains, in order to assess complementary attributes of optogenetic encoding of temporal information. Furthermore, we put a focus on temporal fine structure and the energy

requirements of optogenetic stimulation of the auditory nerve, which are key parameters to be considered in future coding strategies. Finally, we performed a side-by-side comparison of optogenetic and acoustic stimulation employing midbrain activity for estimating the equivalent output dynamics and physiological correlates of stimulus intensity.

Model system and technical validation: The data in this study was obtained from the central nucleus of the inferior colliculus of adult, hearing animals that received an intra-modiolar *AAV-hSyn-CatCh*-eYFP-injection into the left cochlea. The ICC was chosen due to its highly conserved tonotopic organization, which enabled us to simultaneously record multi-units responsive to frequencies across large parts of the gerbils hearing range. Guiding the electrode placement by acoustic stimulation via pure tones enabled comparable recording positions – and thereby averaging of optogenetically evoked responses - across animals. The mean tonotopic slope amounted to 4.23 octaves/mm, which is in good agreement with previous studies, both from our and from other laboratories [14,15,37,42]. Few units (13/534) recorded from ventral electrode contacts also had relatively low characteristic frequencies, indicating that these units might originate from regions outside the ICC [37]. However, as 97.6% of units are expected to have resulted from the ICC, we have not attempted to separate these units when analysing responses to optogenetic stimulation. The transduction rates of SGNs reported in this study (21.6 ± 10.7%) are at the lower edge of the previously reported transduction rates of SGNs upon viral injections in adult gerbils (~20–30%) [22,23]. These studies demonstrated homogeneous transgene expression across different cochlear turns, which we have not analyzed in the current study. However, the fact that we were able to optogenetically drive most ICC multi-units (29.2 ± 2.1 out of 32 recording sites) with stimuli of high energy suggests that transgene expression across the whole cochlea was also achieved in this study.

Response properties for individual light pulses: Individual pulses are defined by their duration and intensity. Varying both of these parameters, we found that pulses as short as 80 μs (when presented at 30 mW), and 0.9 ms (when presented at 2 mW) were sufficient to reliably drive the most sensitive units in the auditory midbrain. Recordings in response to various stimuli at a radiant energy of 2 μJ resulted in activation of 20% of multi-units in response to at least one duration-intensity-combination (32/160 units), demonstrating robust optogenetic neural activation with energy thresholds as low as 2 μJ. The average maximum optogenetic response strength that could be reached with stimuli of up to 10 μJ in this study was a d' of 4.37 ± 0.71. This agrees with previous work, where thresholds between 0.76 and 3.49 μJ and response strengths of 4.27–4.4 d' were reported upon optogenetic fiber stimulation (at pulse durations of 1 ms) [14,15]. Yet, the current study yielded two additional insights towards the understanding and future application of optogenetic SGN stimulation: First, we observed that, at comparable radiant energy, duration-intensity combinations favouring short duration of high intensity were more potent in activating the auditory nerve, at least in a certain range (0.1–3 ms pulse duration at an energy <10 μJ). Second, we found that stimuli that were too short to evoke neural responses when presented individually (termed pulselets) could evoke comparable neural responses when assembled into pulses at pulselet-rates well above the maximum firing rates of SGNs [38] and the expected limitations given by the closing kinetics of *CatCh* [35], the opsin used in this study. Both of these observations have important implications for future coding strategies, which will be discussed below.

Temporal response properties: We characterized the temporal fidelity of auditory midbrain activity in response to optogenetic stimulation of the auditory nerve. Here, we found that the discharge rate of multi-units rose with increasing rates of optogenetic stimulation up to around ~100 Hz and then remained relatively stable up to 500 Hz. The vector strength, a measure of spike timing relative to a stimulation cycle, declined to ~0.4 at repetition rates of 100 Hz, and quickly vanished thereafter in most, but not all multi-units (~25% of multi-units still showed some degree of phase locking up to rates of 250 Hz). This is in good agreement with the gap detection threshold of ~9–10 ms, corresponding to the inter-stimulus interval of a 100 Hz pulse train. The lack of temporal coding in many units beyond these rates might be primarily attributed to the relatively slow off-kinetics of the channelrhodopsin-2 variant *CatCh* [35], rather than to a limitation of multi-neuronal midbrain clusters: First, we have performed parallel recordings in response to stimulation with acoustic click trains in naïve gerbils, which showed a similar increase in discharge rate and decrease in vector strength up to ~100 Hz of stimulation, but showed higher discharge rates and vector strength at higher stimulation rates. Second, the dependence of midbrain activity on the rate of optogenetic stimulation is similar to that found in recordings from individual *CatCh*--transduced SGNs, where a drop in discharge rate in response to stimuli beyond 100 Hz was reported in two thirds of the recorded units, and a drop in vector strength was observed between 100 and 150 Hz in all units [22]. Finally, we note that the majority of neuronal responses presented in this study originate from multiple neurons. We could isolate exemplary single units from the ICC which could follow acoustic and optogenetic stimulation up to 300/150 Hz respectively, and hence demonstrate that individual midbrain neurons can actually encode temporal information well above 100 Hz. Yet, we note that future experiments and analysis will be required to test whether this holds true for the majority of ICC neurons and to investigate the temporal coding properties of larger populations of ICC neurons in greater detail. A second reason limiting the temporal coding might be the intrinsic firing behaviour of *CatCh*-transduced SGNs in response to optical stimulation, where two distinct response types were previously described [22]: While two thirds of SGNs responded with a single action potential and a vector strength of ~0.9 up to stimulation rates of 100 Hz, one third of SGNs responded with multiple action potentials and a vector strength between 0.6 and 0.8. Propagating along the auditory pathway, these response types might result in a synchronized onset of IC activity followed by a less synchronized tail. The relative contributions of single-vs multi-spike origin of a given multi-unit in the ICC might then explain (part of) its temporal precision, and hence – together with the variability in transduction rates of SGNs - explain the large variability of phase locking (Fig. 6D/E) and gap detection thresholds (Fig. 5D). On the other hand, vector strength between optogenetic and acoustic stimulation was comparable up to ~100–150 Hz, although we cannot rule out that this phenomenon only holds for multi-unit rather than single-unit activity. This is in contrast to electrical stimulation, where hyper-precision has been reported on the level of the auditory nerve [43] and brainstem [44], suggesting that oCIs might be able to overcome the hyper-synchronized auditory nerve firing observed with electrical cochlear implants.

Dynamic range estimation: Using a conservative approach, we found a dynamic range of 7.5 dB (mW) which could be encoded by optogenetic activation of the SGNs with trains of 1 ms pulses presented at 100 Hz for 51 ms. This value is in good agreement with values reported for stimulation with single light pulses in previous work employing the same opsin (7.8 dB rel. 1 mW) [14]. At this point, it is important to note that the definition of the dynamic range in response to bionic stimulation of the auditory nerve – which includes both optogenetic, but also electric stimulation – is not straightforward: In theory, the dynamic range can be calculated by the formula $n \times \log10(I_{90}/I_{10})$, where I is the stimulus intensity needed to obtain 90 and 10% of the maximum response, respectively, and n is set to 10 in the case of calculations based on power, while it is set to 20 in the case of calculations based on amplitudes. This calculation becomes intricate when considering which stimulus property should be taken into consideration: pulse intensity (radiant flux) or, multiplied by (constant) pulse duration, pulse energy (radiant energy), resulting in dynamic range estimates differing by a factor of two. The same issue applies to electric stimulation, where either the stimulus intensity (current) or the stimulus charge might be used for dynamic range calculations. In order not to overestimate the potential of optogenetic stimulation, we have previously chosen to calculate the dynamic range in response to optical stimulation based on power, while we opted

for amplitude-based calculations in the case of electrical stimulation, in order to enable literature comparisons [14,17]. Regardless of this issue, the overall dynamic range is probably not the most accurate measure for intensity resolution of bionic auditory nerve stimulation, as it does not contain any information about whether stimuli might be discriminable or not, and hence it does not inform whether the different intensities in between threshold and saturation could even be resolved on a perceptual level (discussed in Ref. [45]). Therefore, in the current study, we now developed an approach to estimate the dynamic range based on a cross-modal comparison of d' values in response to optical stimulation in *CatCh*-injected animals and acoustic stimulation in non-injected, naïve animals. We directly relate the intensities of both modalities, using acoustic stimulation to calibrate the optogenetic counterpart. This way, we estimated the dynamic range in response to optical stimulation equivalent to the dynamic range of acoustic stimulation eliciting similar response strengths. The dynamic range of optogenetic stimulation was equivalent to ~25.6 dB (pe SPL), indicating that we previously indeed underestimated the output dynamic range (7.8 dB rel. 1 mW) [14] available for future oCI coding. This output dynamic range is lower than that of acoustic stimulation most likely because direct optogenetic SGN stimulation bypasses physiological cochlear micromechanics and synaptic sound encoding at the level of inner hair cells. Yet, it substantially exceeds the dynamic range reported for electrical cochlear implants, which is typically below 10 dB (current level) [40,41,46,47]. We note that our assumption of equivalent IC activity levels translating into similar intensity percept awaits future evaluation by psychophysical experiments, which would also allow for a direct comparison to eCI performance. Evaluating threshold, dynamic range, and the maximal neural response strength in the afferent auditory pathway for optogenetic SGN stimulation with ultrafast ChRs and multichannel oCIs are further important objectives for future studies. Using efficient optogenetic SGN stimulation by *CatCh* we provide a first estimate of these values and a framework for the development of coding strategies. We expect this framework to be helpful for devising coding strategies for optogenetic stimulation also when using opsins with improved kinetics such as f-Chrimson [17,20]. These studies will then provide updated estimates for which we expect improved temporal properties of optogenetic sound encoding.

Implications for future coding strategies: When designing future coding strategies, several aspects of optical sound encoding should be considered. The most basic building block of future coding strategies, an individual optical pulse or pulselet, is a fundamental parameter assessed in the present study. We suggest this parameter to range from 0.3 to 1 ms in the case of pulses (Fig. 3B–D, Fig. S6B), and suggest 20 μs as a first estimate for pulselets, as it was shown to efficiently activate *CatCh* in this study. However, given the likely temporal integration of pulselet-mediated photocurrents at high stimulation rates, we expect the range of pulselet durations to be broader. Generally, it seems advisable to balance duration and intensity of single pulses in favor of pulse intensity, minimizing the pulse durations, as pulses of higher intensity – yet similar radiant energy – have been shown to evoke stronger responses (Fig. 3B–D), an aspect of great importance when aiming at minimizing energy requirements for acceptable battery lifetimes in future clinical oCIs. The aspect of energy efficient coding with short pulses is strengthened by the fact that laser diodes are more efficient when being driven with ultrashort pulses of large (tens of mA) current. However, maximizing the ratio of intensity over pulse duration will eventually reach limits: First, the maximum intensity available for optical sound encoding will be limited by the maximum intensity future oCIs can provide (likely up to a few tens of mW). Second, high stimulus intensities might cause phototoxicity, especially when using light in the blue spectral range [48], and hence the operating range of oCIs needs to be limited into a biologically safe range. This could, however, be compensated by lowering the stimulation rates, and hence the radiant energy dose. Also, it should be noted that the ideal ratio of pulse intensity to pulse duration might vary in view of maximizing different aspects of coding. If temporal fidelity is prioritized, one cannot afford to majorly extend pulse durations, as this will decrease the maximum stimulation rate. However, as the strength of neuronal responses scales with energy rather than radiant flux, increasing the radiant energy via prolonged pulse durations when the maximum radiant flux of an oCI emitter has been reached might offer the possibility to further enhance the maximum response strength, and hence the maximum stimulus intensity that can be encoded. Also, it should be kept in mind that not only response strength, but also the spectral selectivity (i.e. the tonotopic range in which SGNs are activated) scales with its energy. This recruitment of additional neurons around the place of stimulation, besides carrying spectral information, is also involved in the encoding of stimulus intensity. While this population code for stimulus intensity offers the possibility for an increase in dynamic range, it also highlights the trade-off between optimal encoding of intensity and spectral information of a stimulus.

The next parameter to be determined in a coding strategy is the repetition rate of individual stimuli, which will convey information about the temporal structure of the encoded sound. As opsins considered for optogenetic hearing restoration still have off-kinetics in the low ms range, sufficient light-off time between individual stimuli is probably needed in order to discriminate these stimuli. Hence, the maximum stimulation rates will be determined by two opsin properties: First, the closing kinetics of the respective opsin used. In our case, this limitation was in the range of ~150 Hz for most units, yet a quarter of all units could keep up with stimulation frequencies of at ~250 Hz. It is important to note that this rate limitation is most likely attributed to the specific kinetics of *CatCh* and can be overcome, as different studies involving faster opsins have demonstrated [17,19,20]. Second, the recovery time for permitting another successful cycle of stimulation requires a sufficiently long dark period and this, together with the stimulus duration, will limit the maximum stimulation rate. The recovery time depends on the cumulative duration of light stimulation in a given cycle. It will be interesting for future studies to parametrize the impact of the precise opsin and stimulus properties on the recovery time of optogenetic stimulation of the auditory nerve and first data on this were recently published [18]. This should also consider the impact of stimulus duty cycle and intensity.

After discussing individual stimuli and their assembly into trains on a single emitter, we next consider the parallel operation of multiple optical stimulation channels in future coding strategies. As discussed above, the encoding of spectral and intensity information will be based on modulations of stimulus intensity and recruitment of neighboring light emitters placed along the tonotopic axis of the cochlea. Hence, optical coding strategies should map independent stimulation channels to perceptually different channels both in terms of spectral and intensity perception and, different from electrical coding, operate them quasi in parallel. Our finding that individual optogenetic stimuli can be assembled from μs-scale, sub-threshold pulselets (Fig. 4) paves the way for "quasi-parallel" stimulation of emitters in the matrix when activating them in a phase-shifted manner. This matrix addressing of the emitters seems mandatory for future coding strategies, as independent electrical addressing of the planned dozens of individual emitters is technically challenging if not impossible. Finally, optical stimulation via μs-scale pulselets enables most efficient operation of semiconductor emitters such as LEDs and laser diodes.

Limitations of the current study: The current study was built on single-channel, fiber-based optogenetic stimulation. Projecting the laser light from the round window along the cochlear modiolus we aimed at broad and robust SGN activation. This facilitated our investigation of stimulus-response-characteristics for individual pulses and pulse trains as required for developing a framework for future oCI coding strategies. Yet, with this configuration we cannot assess the relevant question of overlapping optical stimulation from neighboring emitters, which is an important matter of future investigations. Furthermore, SGN stimulation was limited by the slow off-kinetics of *CatCh*, the opsin used in this

study, as revealed by the analysis of temporal fidelity of IC activity (Fig. 6). ChRs with ultrafast (sub-to low millisecond) off-kinetics such as Chronos [49] or f-Chrimson [20] will enhance the temporal fidelity of optogenetically driven SGN firing [17,19,20]. *CatCh*, on the other hand, confers good light sensitivity to SGNs, which finally determines the threshold of activation and the strength of responses on the level of SGNs. Furthermore, the activation threshold could be lowered – and the dynamic range likely increased – in future studies by decreasing the distance between SGNs and the optical emitter. Indeed, we previously showed that the threshold for neural activation was ~50% lower when using an intra-cochlear, μLED-based oCIs as compared to a laser-coupled optical fiber placed at the round window [15]. A final limitation addresses the specificity of SGN transduction with optogenetic tools in this study. Unlike previous studies that used AAV2/6 [22], using viral vectors of the serotypes PHP.B and PHP.eB, we have sporadically observed individual cochlear hair cells expressing *CatCh*-eYFP despite the use of the synapsin promoter that is not physiologically active in hair cells [36,50]. However, we found few transduced hair cells only in a subset of animals and the strength and latency of neural responses in the ICC were in good agreement with the strength of evoked responses in previous studies [15,22,23] some of which employed deafened animals [15]. Therefore, we argue that the contribution of optogenetically activated hair cells if – at all – is a very minor one, as compared to the contribution of SGNs.

Future perspectives: Despite the recent progress in developing optogenetic hearing restoration, several questions remain to be addressed before potential clinical application. First and foremost, viral transduction of the auditory nerve should be further optimized for efficiency and specificity. While 75–90% of SGNs express opsins in >90% of injected mice [17,20,23,25] and gerbils [23] when animals undergo surgery during the early postnatal period, only ~45% of animals injected at the adult stage showed opsin expression [15,22], which then was restricted to ~20–30% of SGNs [22,23]. In parallel to the optimization of viral expression, studies of biosafety and stability of ChR expression in the cochlea and of the overall biodistribution should be performed in the translational model of the adult gerbil. Investigation of multichannel optogenetic stimulation of the cochlea will be required and should include individual stimuli as presented in the current study, but also focus on intensity and frequency coding. Ideally, these experiments would go alongside behavioural experiments, involving discrimination tasks for psychophysical estimation of the frequency resolution of multi-channel oCIs. Finally, strategies for optogenetic sound encoding remain to be established – a challenge for which we provide a first orientation in the current study. These coding strategies will differ from coding strategies of electrical implants at least with respect to pulse design, stimulation frequencies, and the number of independent stimulation channels that need to be orchestrated.

4. Material/methods

Animals: Data was obtained from 17 Mongolian gerbils (*Meriones unguiculatus*) of either sex between 10 and 56 weeks of age at the start of the experiment. Animals were obtained from the breeding colony at University Medical Center Göttingen and housed in a 12/12 h light-dark-cycle with access to food and water *ad libitum*. All experimental procedures were performed according to German national animal care guidelines and approved by the animal welfare office of the state of Lower Saxony, Germany, as well as the local animal welfare committee of the University Medical Center Göttingen.

Surgical approaches: All surgical procedures were performed on isoflurane-anesthetized animals (3–5% isoflurane at a flow rate of 1 l/min for induction, 0.6–2% isoflurane at a flow rate of 0.4 l/min for maintenance) placed on a heating pad to maintain body temperature. Depth of anesthesia was monitored by the absence of the hind limb withdrawal reflex and adjusted if necessary. Analgesia was achieved by injections of Buprenorphine (0.1 mg/kg bodyweight (BW) s.c.) and Carprofen (5 mg/kg BW s.c.) before surgery. During electrophysiological recordings, Buprenorphine (0.1 mg/kg BW s.c.) was given every 4 h, and Robinul (0.5 mg/kg BW s.c.) was applied to reduce nasopharyngeal secretion if needed.

AAV-injections: Injections of AAV-suspension were performed on adult animals (at least 4 months of age) as described by Wrobel et al. [22]. Briefly, a retro-auricular incision was made behind the left pinna, muscles were replaced, and a bullotomy was performed using a scalpel in order to expose the cochlea. Going through the upper circumference of the round window niche, a small hole was then manually drilled into the base of the modiolus, which houses the spiral ganglion, by using a small dental drill (K-Flex dental file no. 15). Subsequently, 2–3 μl of viral suspension (either AAV-PHP.B [33] [4.6e12 GC/ml] or AAV-PHP.eB [34] [3.8–4.3e12 GC/ml] carrying DNA coding for the calcium translocating channelrhodopsin *CatCh* [35] fused to a reporter protein [enhanced yellow fluorescent protein] under control of the human synapsin promotor) were injected directly into the spiral ganglion via a micropipette (10–20 μm tip diameter) pulled from quartz glass capillaries (P-2000 laser puller, Sutter Instruments) using a pressure microinjector (PLI-100 pico-injector, Harvard Apparatus; 100–125 PSI). The surgical site was then closed by repositioning of connective tissues and suturing the skin. During a recovery period of at least 4 weeks, animals were checked daily and analgesia was administered if necessary (Carprofen, 5 mg/kg BW s.c.; Meloxicam (0.5 mg/kg BW)).

Stimulation: All experiments were performed in a sound-attenuating chamber (Industrial Acoustics; Niederkrüchten, Germany). Stimuli were created in custom-written MATLAB scripts (The MathWorks Inc; Natick, US) actuating a custom-made system based on NI-DAQ-Cards (NI PCI-6229; National Instruments; Austin, US). 30 repetitions have been presented for each stimulus condition. *Acoustic:* Near field acoustic stimulation was performed with a loudspeaker (Scanspeak Ultrasound; Avisoft Bioacoustics; Glienicke, Germany) centered ~30 cm in front of the animal's head. The loudspeaker was calibrated with a 0.25 inch microphone (4039; Brüel & Kjaer; Naerum, Denmark) and corresponding pre- (2670), and measurement-amplifier (2610). *Optogenetic:* For optogenetic stimulation, the cochlea was accessed by the retro-auricular approach described for viral injections (see above). An optical fiber (200 μm diameter, 0.39 NA; Thorlabs; Bergkirchen, Germany) coupled to a 488 nm (LBX-488-100-CSB; Oxxius; Lannion, France; in the case of multi-unit recordings) or a 473 nm (MLLFN-473-100; Changchun New Industry Optoelectronics; China; in the case of oABR recordings) laser was then placed in the cochlea via the round window, facing the cochlear apex and thus illuminating large portions of the cochlea. The output of the optical fiber was measured before each experiment using a power-meter (Solo-2; Gentec-EO; München, Germany) and adjusted to match the reported values ± 10%. Despite this small jitter in absolute laser intensity between animals, the relative intensities used for stimulation within individual animals were always identical (i.e. x % of the maximum power). The temporal structure of stimuli - especially assembled from 20 μs pulses and with repetition rates in the kHz range – was verified using a photodiode connected to an oscilloscope with ~14.3 ns resolution (70 MHz sampling rate; tds2004c; Tektronix Beaverton, US; Fig. S4).

oABR recordings: Optically evoked auditory brainstem responses (oABRs) were recorded as described previously [22]. Briefly, potentials were recorded via two subdermal, low-impedance needle electrodes at the vertex and mastoid bone, while a third needle electrode in the animal's neck was used for active shielding. Potentials were amplified with a custom-made amplifier, digitized at a sampling rate of 50 kHz with a NI-DAQ-Card (National Instruments; Austin, US), and stored on a hard drive. Data were offline filtered between 0.3 and 3 kHz and averaged in response to 1000 repetitions.

Multi-unit recordings: Multi-unit recordings from the ICC have been described in detail before [14–16]. Briefly, an incision was made along the midline in the scalp and the skull was cleaned. A thin layer of UV-glue (Orbi-Bond; Orbis Dental; Münster, Germany) was applied and

cured, before attaching a custom-made metal pin rostral to Bregma using dental cement (Paladur; Kulzer; Hanau, Germany). The animal was stereotactically aligned (less than 100 μm difference between Bregma and Lambda, both on the mediolateral and dorsoventral axis), and a low impedance metal wire was placed between the skull and the cortical surface via a small craniotomy contralateral to the recording site to serve as a reference electrode. Finally, a craniotomy was performed above the visual cortex, which covers the gerbil inferior colliculus, and the dura was removed with a sharp needle. A linear 32-channel electrode array (electrode area 177 μm^2, 1–3 MΩ impedance, 50 μm pitch; Neuronexus, Ann Arbor, US) was then slowly inserted to an initial depth of 3.3 mm (~2 mm lateral (avoiding to injure the cortical vasculature) and as close as possible to the transverse sinus) using a micromanipulator (Luigs & Neumann; Ratingen, Germany). 30 min after insertion, logarithmically spaced pure tones (100 ms duration, 5 ms sine squared ramps for on- and offset; 0.5–32 kHz in quarter octave steps, 10–80 dB SPL in steps of 10 dB SPL) were presented in order to reveal the tonotopic axis of the ICC, and the electrode array position was – if necessary – adjusted to optimally access the ICC and achieve comparable electrode placement across animals. Once the array was positioned, neuronal activity registered from the electrodes was amplified, filtered (0.1–9000 Hz), digitized at a sampling rate of 32 kHz using a Digital Lynx 4S recording system (Neuralynx; Dublin, Ireland) and stored on a hard drive for offline analysis.

Analysis: Multi-unit recordings were filtered using a 4th order Butterworth filter (bandpass, 0.6–6 kHz). A threshold (mean minus three standard deviations of data from the inter-stimulus intervals (i.e. −125 ms to −25 ms before each stimulus onset)) was applied, and crossings of this threshold were defined as neuronal events (a 1 ms refractory period was implemented after each event to avoid over-estimation of responses). For the extraction of exemplary single units, raw data was fed into the template-matching-based, automated spike sorting algorithm *Kilosort3* (*https://github.com/MouseLand/Kilosort*) [51]. After spike sorting, putative units were manually inspected and refined using *Phy* (*https://github.com/cortex-lab/phy*).

Tonotopy: Spike rates during stimulus presentation of 100 ms pure tones (see above) were sorted into a frequency-intensity-matrix. The frequency eliciting responses at the lowest sound pressure level during the period of stimulus presentation was then manually defined as the characteristic frequency [52].

D-prime analysis: To determine potential response windows to various optical stimuli and acoustic clicks, peri-stimulus time histograms (bin size: 0.25 ms) were constructed based on multi-unit activity pooled from all stimuli and electrodes of a given recording paradigm. Response windows were defined as (at least) two subsequent bins exceeding a threshold of the mean plus three standard deviations during the 10 ms before stimulus onset, and ranged from 2.25 to 38.25 ms for optogenetic experiments covering a wide range of pulse durations and intensities (Fig. 2), 1.75–27.75 ms for experiments using energy equivalent pulses (Fig. 3), 2–26 ms for experiments using high frequency stimulation (Fig. 4), 2–114.5 ms for experiments using 101 ms pulse trains of 1 ms pulses presented at varying stimulation frequencies (Fig. 6), and 1.75–63.25/2–68.5/1.75–70.5 ms for experiments using 51 ms pulse trains of 1 ms pulses presented at varying intensity at a stimulation rate of 50/100/200 Hz (Fig. 7). Based on these findings, windows for analysis were chosen from 0 to 40, 0–30, 0–25, 0–115, and 0–65/70/75 ms, respectively. In the case of acoustic stimulation, response windows were found to range from 4.5 to 121.25 ms for experiments using 101 ms click trains of 0.3 ms clicks repeated at varying stimulation frequencies (Fig. 6) and 4.5–92.5 ms for experiments using 51 ms click trains of 0.3 ms clicks of varying intensity at a stimulation rate of 100 Hz (Fig. 7), and analysis windows were accordingly set from 0 to 125 and 0–85/95/100 ms. The strength of neuronal responses was then calculated as the discrimination index (d') based on spike rate distributions observed in the previously defined response windows versus the spike rate distributions observed in a time window of the same duration just before stimulus onset. For the discrimination of two stimuli (or one stimulus and one baseline condition) presented 30 times each, this resulted in a total of 60 measures of spike counts. From these spike counts, we generated an empirical receiver operating curve (ROC) and calculated the resulting area under the curve (AUC). The d' was then calculated as the square root of two multiplied by the AUC transformed with the inverse normal cumulative distribution function [53,54]. In the case of non-overlapping distributions, which would result in an AUC of 0 or 1, both true and false positive rates were corrected by the 1/2 N procedure for the correction of extreme values [53]. Given the equal number of trials for both conditions, this correction corresponds to $1/(2 \times N)^2$, resulting in AUCs of 0.0003 and 0.9997, respectively, and a maximum possible d' value of |4.88|.

Generally, we observed that baseline activity was higher in optogenetically stimulated animals as compared to acoustically stimulated animals (42.3 ± 14.5 Hz vs 28.2 ± 7.8 Hz, $p = 1.4 \times 10^{-82}$, Wilcoxon rank sum test, n = 928/384 recordings each). While we do not have a definite answer on why this might be the case and reasons might be manifold (e.g. chronic changes upon cochlear surgery for virus injection, acute changes upon fiber implantation, unilateral fiber stimulation vs open field acoustic stimulation, …), we note that we have performed all analysis based on d' values rather than absolute spike rates, i.e. based on relative changes in response strength rather than absolute activity levels, in order to compensate for uneven baseline activity. Responding units were defined as units showing a d' of 1 or higher in response to at least one stimulus of a given stimulus set.

Gap detection: To determine gap detection thresholds, we first assembled peri-stimulus time histograms (PSTHs) for each stimulus by summing up the detected spikes of each stimulus repetition (bin size: 2 ms). Using a two-tailed *t*-test, spike rates during the gap window (and an additional delay of 3 ms to compensate for the response offset) of stimuli with various gap durations were compared to spike rates during a reference time window of 10 ms duration which was taken from the same starting point in stimuli without a gap (no-gap condition). The first gap stimulus with a significant reduction in spike rate was defined as the gap detection threshold, provided that all the gaps of longer duration also showed this reduction (although one non-significant exception was allowed in the series) and provided that the first significant reduction in spike rate was found before the largest presented gap.

Temporal response properties: For the analysis of temporal response properties, only responsive units (i.e. units reaching a d' of 1 or higher in response to a single pulse of 1 ms duration and 32 mW intensity) were considered. The discharge rate of these units was defined as the number of spikes in the determined response window (see above) divided by the duration of this response window. Vector strength was calculated as

$$VS = \frac{\sqrt{\left[\sum_{i=1}^{n} cos\ \theta i\right]^2 + \left[\sum_{i=1}^{n} sin\ \theta i\right]^2}}{n}$$

with θ depicting the phase of a multi-unit spike in a given stimulus cycle (i.e. time from stimulus onset to subsequent stimulus onset) [22,55]. To avoid a bias in vector strength caused by the onset response of each stimulus train, only responses during the second half of the stimulus train, i.e. earliest 50 ms after stimulus onset, were considered. The Rayleigh test was used to probe the significance of the vector strength, and vector strength below $L < 13.8$ (i.e. $p > 0.001$) was considered insignificant and set to 0.

Dynamic range: The dynamic range of optical SGN stimulation was calculated as the difference in stimulus intensities that increased firing rates of IC multi-units from 10% above baseline to 10% below saturation. To this end, baseline activity was defined as the average firing rate in response to the three lowest stimulus intensities, whereas saturation was defined as the firing rate in response to the three highest stimulus intensities. Only units with at least a two-fold increase in baseline firing rates were considered for this analysis. To relate the dynamic range of

optogenetically to acoustically evoked IC activity, a cumulative discrimination index (d') was calculated based on the multi-unit firing rates in response to the respective stimulus modality [14]. First, for each animal, responses to either optogenetic or acoustic stimuli were sorted according to increasing intensity. The d' value was then calculated for each pair of neighboring stimulus intensities, i.e. to discriminate between the responses to a given stimulus and the responses to the stimulus with the next higher intensity. For a given comparison between successive stimulus intensities presented 30 times each, this resulted in 60 measures of spike counts. From these spike counts, we generated an empirical receiver operating curve (ROC) and calculated the resulting area under the curve (AUC). The d' was then calculated as the square root of two multiplied by the AUC transformed with the inverse normal cumulative distribution function [53]. Extreme values of the AUC (0 and 1) were set to 0.0003 and 0.9997, respectively. This results in a maximum possible d' value of |4.88|. These d' values of successive stimulus intensities were then summed up, in order to obtain the cumulative discrimination index [54,56]. Finally, we related stimulus intensities which elicited neuronal responses of a given strength – either in response to acoustic or in response to optogenetic stimuli – to each other in order to compare the absolute stimulus intensities evoking neural responses of similar strength across different stimulus modalities.

Statistics: Throughout the manuscript, statistical differences across two groups were quantified using a *t*-test (in case of normally distributed data) or a Wilcoxon rank sum test (for non-normal distributions), as indicated, with Bonferroni-corrected alpha levels if multiple conditions were compared. Normality was tested using the Jarque-Bera-test. Statistical differences across groups were quantified using analysis of variance (ANOVA) and Tukey's multiple comparison tests (if statistical differences were indicated by ANOVA). Data is reported as mean ± standard deviation, unless indicated otherwise.

Histology: Upon completion of the experiment, deeply anesthetized animals were sacrificed and cochleae were explanted, dissected, perfused with 4% Paraformaldehyde (PFA) in phosphate buffered saline (PBS) via the round window, and fixed in the same solution for 30–60 min. Cochleae were then decalcified in ethylenediaminetetraacetic acid (EDTA; 0.12 M) for 5–7 days, cryoprotected in 25% sucrose in PBS for at least 24 h, and sectioned into 16 μm thick slices using a cryostate (CM3050S, Leica, Germany). Sections were washed for 3 times for 5 min each with PBS and unspecific binding sites were blocked for 1 h at room temperature in Goat Serum Dilution Buffer (GSDB; 10 ml goat serum, 1,8 ml 10% triton X-100, 5 ml 240 mM phosphate buffer, and 6.75 ml 4 M NaCl in 36.45 ml H_2O). Subsequently, slices were incubated with guinea pig anti-parvalbumin (1:300; 195004, Synaptic Systems), and chicken anti-GFP (1:500, ab13970, Abcam) primary antibodies in GDSB, either for 2 h at room temperature or overnight at 4 °C. Afterwards sections were washed 3 times for 5 min with wash buffer (83 ml 240 mM phosphate buffer, 30 ml 10% triton X-100, 112.5 ml 4 M NaCl and 1L double-distillated H_2O), before being incubated with secondary antibodies (goat anti-guinea pig (1:200; Alexa Fluor 568, A11075, Invitrogen), and goat anti-chicken (1:200, Alexa Fluor 488, A11039, Invitrogen)) in GSDB for 1 h in darkness. Prior to embedding, slices were washed 3 × 5 min in wash buffer and then for 5 min in PBS. After embedding of the slices, images were taken with a LSM510 (Zeiss, 467 Jena, Germany) microscope and analyzed with ImageJ/Fiji [57].

CRediT authorship contribution statement

Maria Michael: Investigation, Validation, Data curation, Writing – original draft, Writing – review & editing. **Bettina Julia Wolf:** Investigation, (injections), Validation, Writing – review & editing. **Astrid Klinge-Strahl:** Investigation, (gap detection), Data curation, (gap detection), Formal analysis, (gap detection), Validation, Writing – review & editing. **Marcus Jeschke:** Conceptualization, Methodology, Software, Validation, Writing – review & editing, Supervision. **Tobias Moser:** Conceptualization, Validation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Alexander Dieter:** Conceptualization, Methodology, Software, Investigation, (injections), Formal analysis, Validation, Writing – original draft, Writing – review & editing, Visualization, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: TM is co-founder of the OptoGenTech Company. Remaining authors declare no conflict of interest.

Acknowledgements

The authors gratefully acknowledge Dr. Vladan Rankovic and Dr. Kathrin Kusch for production of plasmids and virus used in this study, Daniela Gerke for assistance with virus production and histology, and Carlotta Geyer for help with histology. We further gratefully acknowledge Dr. Jakob Neef and Dr. Antoine Huet for assistance with confocal microscopy and image analysis. We thank Gerhard Hoch for excellent technical support, and Patricia Räke-Kügler for excellent administrative support. This work was funded by the European Research Council through the Advanced Grant 'OptoHear" to TM under the European Union's Horizon 2020 Research and Innovation program (grant agreement No. 670759) and by the German Research Foundation through the Cluster of Excellence (EXC2067) Multiscale Bioimaging to TM. This research was further supported by Fondation Pour l'Audition (FPA RD-2020-10) to TM. MM was supported by a scholarship of the Göttingen Promotionskolleg für Medizinstudierende, funded by the Else Kröner-Fresenius-Stiftung (Promotionskolleg für Epigenomik und Genomdynamik, 2017, Promotionskolleg 04). AD is a fellow of the German Academic Scholarship Foundation (Studienstiftung des Deutschen Volkes).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brs.2023.09.018.

References

[1] Middlebrooks JC, Bierer JA, Snyder RL. Cochlear implants: the view from the brain. Curr Opin Neurobiol 2005;15:488–93.
[2] Lenarz T. Cochlear implant - state of the art. GMS Curr Top Otorhinolaryngol, Head Neck Surg 2017;16:Doc04.
[3] Kleinlogel S, Vogl C, Jeschke M, Neef J, Moser T. Emerging approaches for restoration of hearing and vision. Physiol Rev 2020;100:1467–525.
[4] Ramsden RT. History of cochlear implantation. Cochlear Implants Int 2013;14 (Suppl 4):S3–5.
[5] Kral A, Hartmann R, Mortazavi D, Klinke R. Spatial resolution of cochlear implants: the electrical field and excitation of auditory afferents. Hear Res 1998;121:11–28.
[6] Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. J Acoust Soc Am 2001;110:1150–63.
[7] Garnham C, O'Driscoll M, Ramsden And R, Saeed S. Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets. Ear Hear 2002;23:540–52.
[8] Fishman KE, Shannon RV, Slattery WH. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. J. Speech Lang. Hear. Res. JSLHR 1997;40:1201–15.
[9] Wolf BJ, et al. Is there an unmet medical need for improved hearing restoration? EMBO Mol Med 2022;14:e15798.
[10] Hunniford V, et al. Patient perspectives on the need for improved hearing rehabilitation: a qualitative survey study of German cochlear implant users. Front Neurosci 2023;17.
[11] Dieter A, Keppeler D, Moser T. Towards the optical cochlear implant: optogenetic approaches for hearing restoration. EMBO Mol Med 2020;12:e11618.
[12] Hernandez VH, et al. Optogenetic stimulation of the auditory pathway. J Clin Invest 2014;124:1114–29.
[13] Moser T, Dieter A. Towards optogenetic approaches for hearing restoration. Biochem Biophys Res Commun 2020;527:337–42.
[14] Dieter A, Duque-Afonso CJ, Rankovic V, Jeschke M, Moser T. Near physiological spectral selectivity of cochlear optogenetics. Nat Commun 2019;10:1962.

[15] Dieter A, et al. µLED-based optical cochlear implants for spectrally selective activation of the auditory nerve. EMBO Mol Med 2020;12:e12387.
[16] Keppeler D, et al. Multichannel optogenetic stimulation of the auditory pathway using microfabricated LED cochlear implants in rodents. Sci Transl Med 2020;12.
[17] Bali B, et al. Utility of red-light ultrafast optogenetic stimulation of the auditory pathway. EMBO Mol Med 2021. https://doi.org/10.15252/emmm.202013391.
[18] Mittring A, Moser T, Huet AT. Graded optogenetic activation of the auditory pathway for hearing restoration. Brain Stimul 2023;16:466–83.
[19] Keppeler D, et al. Ultrafast optogenetic stimulation of the auditory pathway by targeting-optimized Chronos. EMBO J 2018;37.
[20] Mager T, et al. High frequency neural spiking and auditory signaling by ultrafast red-shifted optogenetics. Nat Commun 2018;9:1750.
[21] Duarte MJ, et al. Ancestral adeno-associated virus vector delivery of opsins to spiral ganglion neurons: implications for optogenetic cochlear implants. Mol. Ther. J. Am. Soc. Gene Ther. 2018;26:1931–9.
[22] Wrobel C, et al. Optogenetic stimulation of cochlear neurons activates the auditory pathway and restores auditory-driven behavior in deaf adult gerbils. Sci Transl Med 2018;10.
[23] Huet AT, Dombrowski T, Rankovic V, Thirumalai A, Moser T. Developing fast, red-light optogenetic stimulation of spiral ganglion neurons for future optical cochlear implants. Front Mol Neurosci 2021;14:635897.
[24] Zerche M, Wrobel C, Kusch K, Moser T, Mager T. Channelrhodopsin fluorescent tag replacement for clinical translation of optogenetic hearing restoration. Mol. Ther. - Methods Clin. Dev. 2023;29:202–12.
[25] Keppeler D, et al. Multiscale photonic imaging of the native and implanted cochlea. Proc Natl Acad Sci USA 2021;118:e2014472118.
[26] Khurana L, Keppeler D, Jablonski L, Moser T. Model-based prediction of optogenetic sound encoding in the human cochlea by future optical cochlear implants. Comput Struct Biotechnol J 2022;20:3621–9.
[27] Klein E, Gossler C, Paul O, Ruther P. High-density µled-based optical cochlear implant with improved thermomechanical behavior. Front Neurosci 2018;12:659.
[28] Goßler C, et al. GaN-based micro-LED arrays on flexible substrates for optical cochlear implants. J Phys Appl Phys 2014;47:205401.
[29] Schwaerzle M, Nehlich J, Ayub S, Paul O, Ruther P. Led-based optical cochlear implant on highly flexible triple layer polyimide substrates. In: 2016 *IEEE 29th international Conference on micro electro mechanical systems (MEMS)* 395–398. IEEE; 2016. https://doi.org/10.1109/MEMSYS.2016.7421644.
[30] Jablonski L, et al. Hearing restoration by a low-weight power-efficient multichannel optogenetic cochlear implant system. 2020. http://biorxiv.org/lookup/doi/10.1101/2020.05.25.114868. 10.1101/2020.05.25.114868.
[31] Gordon-Salant S. Speech perception and auditory temporal processing performance by older listeners: implications for real-world communication. Semin Hear 2006;27:264–8.
[32] He S, et al. Gap detection measured with electrically evoked auditory event–related potentials and speech-perception abilities in children with auditory neuropathy spectrum disorder. Ear Hear 2013;34:733–44.
[33] Deverman BE, et al. Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. Nat Biotechnol 2016;34:204–9.
[34] Chan KY, et al. Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. Nat Neurosci 2017;20:1172–9.
[35] Kleinlogel S, et al. Ultra light-sensitive and fast neuronal activation with the Ca^{2}+-permeable channelrhodopsin CatCh. Nat Neurosci 2011;14:513–8.
[36] Safieddine S, Wenthold RJ. SNARE complex at the ribbon synapses of cochlear hair cells: analysis of synaptic vesicle- and synaptic membrane-associated proteins. Eur J Neurosci 1999;11:803–12.
[37] Schnupp JWH, Garcia-Lazaro JA, Lesica NA. Periodotopy in the gerbil inferior colliculus: local clustering rather than a gradient map. Front Neural Circ 2015;9.
[38] Liberman MC. Auditory-nerve response from cats raised in a low-noise chamber. J Acoust Soc Am 1978;63:442–55.
[39] Pfingst BE, Holloway LA, Razzaque SA. Effects of pulse separation on detection thresholds for electrical stimulation of the human cochlea. Hear Res 1996;98:77–92.
[40] Allitt BJ, et al. Midbrain responses to micro-stimulation of the cochlea using high density thin-film arrays. Hear Res 2012;287:30–42.
[41] Allitt BJ, Harris AR, Morgan SJ, Clark GM, Paolini AG. Thin-film micro-electrode stimulation of the cochlea in rats exposed to aminoglycoside induced hearing loss. Hear Res 2016;331:13–26.
[42] Harris DM, Shannon RV, Snyder R, Carney E. Multi-unit mapping of acoustic stimuli in gerbil inferior colliculus. Hear Res 1997;108:145–56.
[43] Hartmann R, Klinke R. Impulse patterns of auditory nerve fibres to extra-and intracochlear electrical stimulation. Acta Otolaryngol 1990;109:128–34.
[44] Müller M, et al. Temporal hyper-precision of brainstem neurons alters spatial sensitivity of binaural auditory processing with cochlear implants. Front Neurosci 2023;16:1021541.
[45] Jeschke M, Moser T. Considering optogenetic stimulation for cochlear implants. Hear Res 2015;322:224–34.
[46] Rubinstein JT. How cochlear implants encode speech. Curr Opin Otolaryngol Head Neck Surg 2004;12:444–8.
[47] Zeng F-G. Trends in cochlear implants. Trends Amplif 2004;8:1–34.
[48] Commission E. Directive 2006/25/EC of the European Parliament and of the Council (artificial optical radiation). Off J Eur Union 2006;114:38–59.
[49] Klapoetke NC, et al. Independent optical excitation of distinct neural populations. Nat Methods 2014;11:338–46.
[50] Nouvian R, et al. Exocytosis at the hair cell ribbon synapse apparently operates without neuronal SNARE proteins. Nat Neurosci 2011;14:411–3.
[51] Pachitariu M, Sridhar S, Stringer C. Solving the spike sorting problem with Kilosort. 2023. http://biorxiv.org/lookup/doi/10.1101/2023.01.07.523036. 10.1101/2023.01.07.523036.
[52] Egorova M, Ehret G, Vartanian I, Esser KH. Frequency response areas of neurons in the mouse inferior colliculus. I. Threshold and tuning characteristics. Exp Brain Res 2001;140:145–61.
[53] Macmillan NA, Creelman D. Detection theory: a user's guide by neil A. Psychology Press; 2004. Macmillan.
[54] Middlebrooks JC, Snyder RL. Auditory prosthesis with a penetrating nerve array. J. Assoc. Res. Otolaryngol. JARO 2007;8:258–79.
[55] Goldberg JM, Brown PB. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. J Neurophysiol 1969;32:613–36.
[56] Dieter A, Duque-Afonso CJ, Rankovic V, Jeschke M, Moser T. Near physiological spectral selectivity of cochlear optogenetics. Nat Commun 2019;10:1962.
[57] Schindelin J, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;9:676–82.