

18. Schmid-Hempel, P. *Parasites in Social Insects* (Princeton Univ. Press, Princeton, 1998).
19. Shykoff, J. A. & Schmid-Hempel, P. Parasites delay worker reproduction in bumblebees: consequences for eusociality. *Behav. Ecol.* **2**, 242–248 (1991).
20. Müller, C. B. & Schmid-Hempel, P. Variation in worker mortality and reproductive performance in the bumble bee, *B. lucorum*. *Funct. Ecol.* **6**, 48–56 (1992).
21. Müller, C. B. & Schmid-Hempel, P. Correlates of reproductive success among field colonies of *Bombus lucorum* L.: the importance of growth and parasites. *Ecol. Entomol.* **17**, 343–353 (1993).
22. Liersch, S. & Schmid-Hempel, P. Genetic variability within social insect colonies reduces parasite load. *Proc. R. Soc. Lond. B* **265**, 221–225 (1998).
23. Walker, W. F. Sperm utilization strategies in non-social insects. *Am. Nat.* **115**, 780–799 (1980).
24. Madsen, T., Shine, R., Loman, J. & Hakansson, T. Why do female adders copulate so frequently? *Nature* **355**, 440–441 (1992).

**Acknowledgements.** We thank K. Boomsma, M. Brown, F. Fischer, B. Imhoof, S. Koulianos, N. Krüger, E. Magro, E. Meier, R. E. Page Jr, F. Ratnieks, R. Schmid-Hempel and H. Schwarz for help and comments, and the local communities and forestry commissions for use of their land. This work was supported by grants to P.S.-H. from the Swiss National Science Foundation and the Swiss Office of Science and Technology (within a European TMR network).

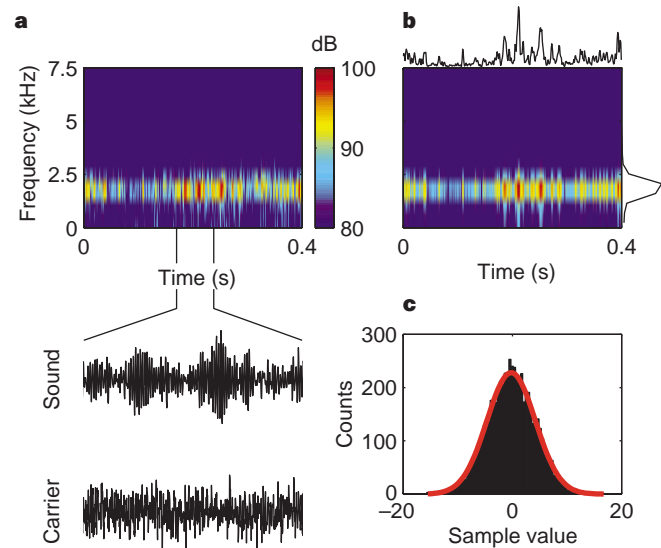
Correspondence and requests for materials should be addressed to P.S.-H. (e-mail: psh@eco.unmw.ethz.ch).

## Responses of auditory-cortex neurons to structural features of natural sounds

Israel Nelken, Yaron Rotman & Omer Bar Yosef

Department of Physiology, Hadassah Medical School, Ein Karem, PO Box 12272, Jerusalem 91120, Israel

Sound-processing strategies that use the highly non-random structure of natural sounds may confer evolutionary advantage to many species. Auditory processing of natural sounds has been studied almost exclusively in the context of species-specific vocalizations<sup>1–4</sup>, although these form only a small part of the acoustic biotope<sup>5</sup>. To study the relationships between properties of natural soundscapes and neuronal processing mechanisms in the auditory system, we analysed sound from a range of different environments. Here we show that for many non-animal sounds

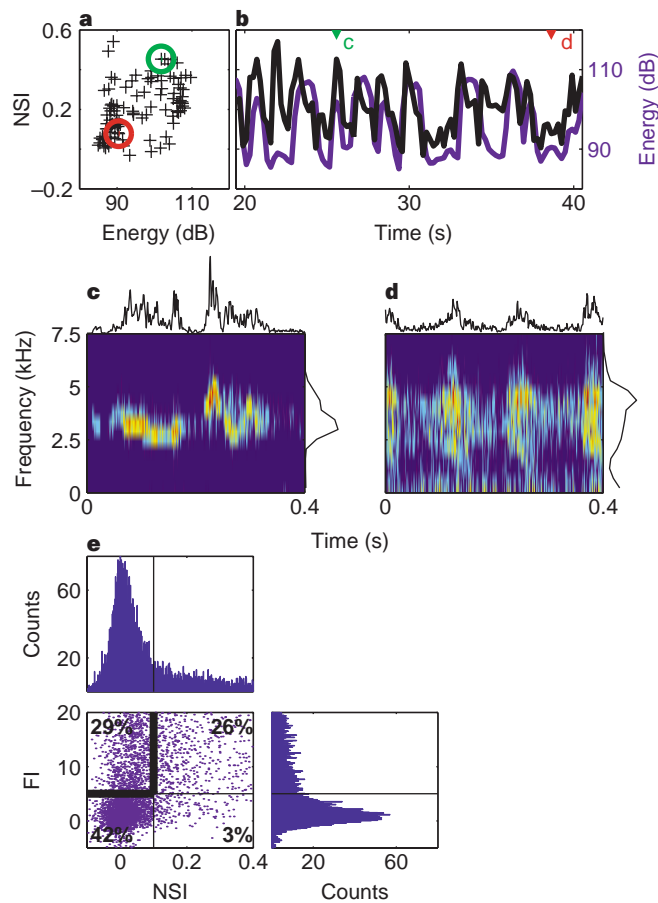


**Figure 1** An example of the separation of a sound into a separable spectrogram and a carrier. **a**, The spectrogram together with the temporal waveform of a segment of the sound and the demodulated carrier. **b**, The separable approximation. The temporal and spectral components are shown at the margins. Same scale as in **a**. **c**, The amplitude distribution of the carrier, with a fit to the gaussian distribution ( $\chi^2 = 74$ , d.f. = 75, not significant). The amplitude distribution of the original sound segment is clearly non-gaussian ( $\chi^2 = 222$ , d.f. = 75,  $P = 0$ ).

and background mixtures of animal sounds, energy in different frequency bands is coherently modulated. Co-modulation of different frequency bands in background noise facilitates the detection of tones in noise by humans, a phenomenon known as co-modulation masking release (CMR)<sup>6,7</sup>. We show that co-modulation also improves the ability of auditory-cortex neurons to detect tones in noise, and we propose that this property of auditory neurons may underlie behavioural CMR. This correspondence may represent an adaptation of the auditory system for the use of an attribute of natural sounds to facilitate real-world processing tasks.

We analysed soundscapes (long recordings of mixtures of animal vocalizations and non-animal sounds) and vocalizations of single animals (for example, frogs, birds, cats and dogs). More than 1300 s from 25 sources were analysed. Following previous work on natural images<sup>8</sup>, the sounds were decomposed into an envelope modulating a gaussian carrier. The spectrogram was used as a spectrotemporal envelope. The parameters of the spectral analysis were fixed at values that resulted in the carrier having gaussian statistics for most sound segments.

Figure 1 shows the spectrogram of a segment from a chorus of crows, which is clearly dominated by a fluctuating horizontal stripe. This structure indicates that the spectrogram is approximately separable: it is well approximated by a product of a function of



**Figure 2** Statistics of soundscapes. **a**, A scatter plot of NSI against energy of 0.4-s sound segments with 0.2-s overlap ( $R = 0.44$ ,  $P < 0.01$ ). **b**, NSI (black line) and energy (purple line) as functions of time (same segment). **c**, A spectrogram of a non-separable, fluctuating segment (NSI = 0.45, FI = 28.6). **d**, A spectrogram of a separable, fluctuating segment (NSI = 0.08, FI = 28.0). The location of the segments in **c** and **d** are shown in **b** by arrowheads, and in **a** by circles. **e**, A scatter plot of NSI against FI for 0.4-s segments covering the whole data set. Only 64% of the segments appear in this plot; most of the others have larger NSI and/or FI values.

frequency alone and a function of time alone. We developed a non-separability index (NSI) to quantify the amount of discrepancy between the spectrogram and its separable approximation (the most similar separable function of time and frequency). In Fig. 1, the spectrogram is essentially equivalent to its approximation (NSI = 0.01; rather small NSI, see Fig. 2e). Segments with a gaussian carrier and separable spectrogram are approximately gaussian processes multiplied by a temporal envelope. We developed a fluctuation index (FI) to quantify the excess fluctuations of the temporal envelope relative to the expected amount for an unmodulated gaussian process with the same power spectrum. The sound segment in Fig. 1 has a significant excess of temporal fluctuations (FI = 17.1). Separable sounds with high excess temporal fluctuations have the important property that energy in different frequency bands is modulated coherently in time.

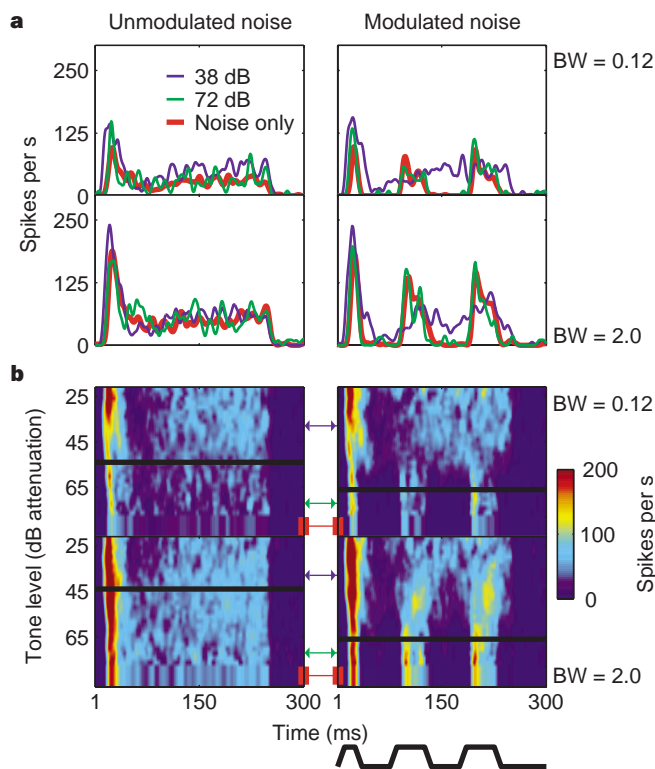
Many animal vocalizations contain frequency-modulated components, in which frequency and time are not separable. In contrast, mixtures of vocalizations, and non-animal sounds, tend to be separable. Figure 2a–d shows the analysis of a soundscape that contained vocalizations of a dominant bird over a background of other animals. The NSI fluctuates over time, and tends to be large when the total energy of the segment is large, corresponding to a call of the dominant bird. Thus, in this case, background sounds tend to be separable. To calibrate the scale of the NSI and FI, Fig. 2e shows their common distribution. The marginal distributions of NSI and FI have a peak near 0, consisting of those sounds that are separable (for the NSI) or have nonsignificant temporal fluctuation (for the FI). Based on these histograms, we considered segments with NSI < 0.1 to be separable and segments with FI < 5 to have nonsignificant envelope fluctuations. Approximately 29% of the

sounds are separable (NSI < 0.1), and also show larger than expected envelope fluctuations (FI > 5).

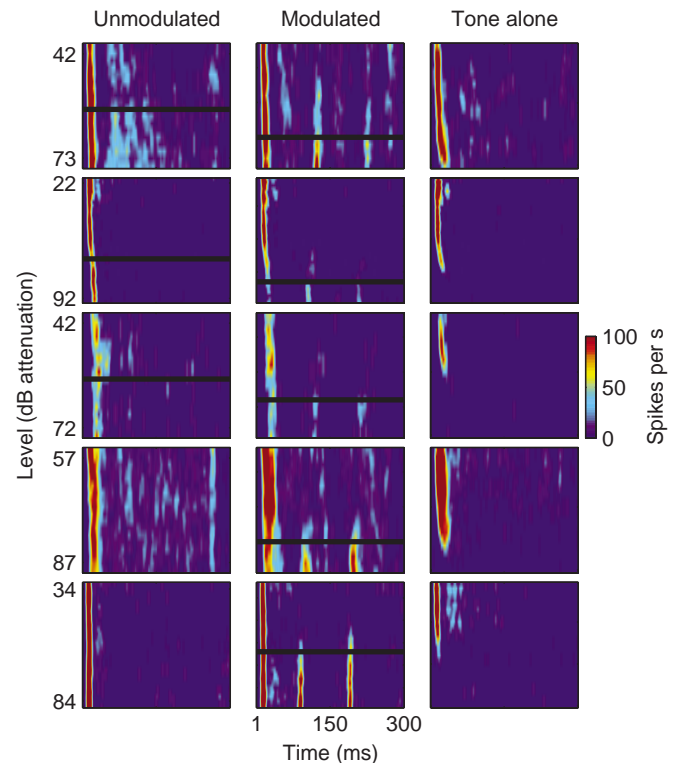
Mixtures of calls are probably separable because of a central limit theorem: when many calls are added together at random times, the result is a stationary gaussian process with the average power spectrum of the original calls. In simulations, five simultaneous calls were sufficient to reach the convergence criterion (NSI < 0.1). However, the energy fluctuations of different frequency bands became decorrelated and FI at convergence was small. Two possible factors may increase co-modulation. First, animal calls sometimes tend to form clusters. This effect was observed occasionally in the data set. Second, sounds acquire slow (< 50 Hz), large envelope fluctuations during atmospheric propagation because of microturbulence<sup>9</sup>.

Thus, separable sounds with significant envelope fluctuations are rather common in the physical world. Because such separable backgrounds can elicit CMR, our results indicate that CMR is an adaptation of the human auditory system for detecting signals over naturally occurring separable backgrounds. We propose that other animals might also profit from the use of the same regularity in the structure of natural sounds. Indeed, it has been shown behaviourally that some birds show CMR<sup>10</sup>, although a physiological study in cochlear nucleus of cats did not find strong effects<sup>11</sup>.

To test this hypothesis further, we searched for neuronal correlates of CMR in primary auditory cortex (AI) of cats (Figs 3, 4). In response to modulated noise bands (MNBs), most of the neurons (35/56, 63%) modulated their firing rates coherently with the temporal envelope (a phenomenon described previously<sup>12</sup> and referred to here as ‘envelope locking’) (Fig. 3a, modulated noise, red lines). In contrast, responses to unmodulated noise bands



**Figure 3** Responses to noise alone and to tone plus noise. **a**, Responses of a neuron to noise alone (red) and tone plus noise at two tone levels (purple and green). **b**, Responses to tone plus noise combinations at many levels. The tone levels used in **a** are marked by arrows (all levels are given in dB attenuation). The thick black lines are tone thresholds in noise. Thresholds: 54 dB (unmodulated), 66 dB (modulated) at 0.125 kHz; 44 dB (unmodulated), 66 dB (modulated) at 2 kHz. BW, bandwidth.



**Figure 4** Additional examples of CMR. Each line shows the responses of one neuron to tone plus unmodulated noise (left), tone plus modulated noise (centre), and tone alone (right). Changes in envelope locking sometimes occur below the neuron’s threshold in silence. From top to bottom: BF = 10.5 kHz, noise bandwidth (BW) = 10 kHz, threshold difference ( $\Delta\theta$ ) = 8 dB; BF = 10.5 kHz, BW = 4.25 kHz,  $\Delta\theta$  = 14 dB; BF = 3.4 kHz, BW = 1.7 kHz,  $\Delta\theta$  = 6 dB; BF = 7.0 kHz, BW = 3.5 kHz,  $\Delta\theta$  > 30 dB; BF = 7.8 kHz, BW = 3.9 kHz,  $\Delta\theta$  > 34 dB. In the last two cases, threshold in unmodulated noise was not reached.

(UNBs) were either pure onset (23/35) or onset followed by unpatterned, sustained increases in firing rates (12/35; Fig. 3a, unmodulated noise, red lines).

We found that in a substantial fraction of the neurons that locked to the temporal envelope (25/35, 71%), the envelope locking was labile and degraded easily when a tone was added to the noise. In Fig. 3a, at the lower tone level (green lines), the response to the tone plus noise combination is essentially identical to the response to the noise alone. In contrast, at the higher tone level (blue lines), the temporal pattern of the response is modified. In responses to UNBs, sustained firing rates are increased, the onset response is larger and its latency is shorter. Similar changes in onset response also occur in responses to MNBs; however, in this case there is also a strong decrease in envelope locking, which is replaced by a more sustained firing pattern. Thus, if the shape of the peristimulus time histogram (PSTH) is used to distinguish between the noise alone and the tone plus noise cases, the detection threshold should occur between the two sound levels in all panels of Fig. 3a. To find the masked threshold, responses at many tone levels are presented as response planes, with firing rate displayed in colour as a function of tone level and time (Fig. 3b). The suppression of envelope locking caused a substantial reduction in detection thresholds of tones in MNBs compared with UNBs, and this reduction increased with bandwidth. Additional examples of the same phenomenon are shown in Fig. 4: neuronal responses lock to the envelope of the MNBs at low tone levels, at which the response to the tone plus noise combination is essentially identical to the responses to the noise alone (as in Fig. 3a at the low tone level). This envelope locking disappears at higher tone levels, giving a qualitative (as well as quantitative) signal for the presence of the added tone.

Envelope locking often increased with bandwidth when measured by locked rate, at least for the narrower bandwidths (23/35, 66%; see refs 13, 14). Thus, the population signal for release from masking in MNBs increased with bandwidth. Population modulated–unmodulated threshold differences, computed from the average response planes of the whole population reported here, were about 10 dB at the narrowest noise bandwidth tested (BF/32 of each neuron, where BF is the best frequency), and over 30 dB at the largest (BF/2 of each neuron). The decrease in threshold in MNBs as bandwidth was increased was over 20 dB in the population signal; using similar stimuli, corresponding values in humans are 10–15 dB (ref. 15). The decrease in thresholds as noise bandwidth is increased beyond the peripheral filters is the defining characteristic of true CMR. The degradation of envelope locking depended on time: later response components were often degraded more easily (for example, Fig. 4, top three examples); indeed, in human psychophysical experiments, sounds must be rather long (>200 ms) to elicit maximal CMR. Thus, neuronal responses and human psychophysics followed similar rules<sup>6,7</sup>. With the cautionary note that the physiology of anaesthetized cats has been compared with the behaviour of humans, we speculate that this neuronal mechanism underlies CMR.

It is not clear whether the effects shown here are generated at the level of the primary auditory cortex. One argument for a cortical origin of these responses is the low modulation frequencies that are required to elicit CMR (above 50 Hz, true CMR is minimal). These rates correspond better to cortical modulation sensitivity than to modulation sensitivity of inferior colliculus neurons, which may follow amplitude modulation up to hundreds of Hz. The lability of the late envelope-following responses probably results from the sluggishness of cortical neurons.

We suggest that these results are a special case of a general correspondence principle for auditory physiology: regularities in the auditory biotope should be manifested in response properties of auditory neurons. Such correspondences have been suggested previously for species-specific vocalizations<sup>4</sup>; we have demonstrated here that this principle may cover a much larger class of auditory phenomena. □

## Methods

**Analysis of natural sounds.** Recordings are the property of the Library of Natural Sounds, Cornell Laboratory of Ornithology, Ithaca, New York. The sounds were copied from digital audio tape to disk files. Sampling rate was reduced from 44.1 to 14.7 kHz, because the sounds did not contain appreciable energy above 7 MHz. Multiple 32-point fast Fourier transforms (FFTs) were computed, shifting the window by one sample between successive computations. The spectrogram was formed by taking the squared absolute value of the samples at the output of each FFT filter and smoothing them with a two-bin window in frequency. The carrier was computed by dividing the outputs of the FFT filters by the square root of the smoothed spectrogram; the result was transformed back into the time domain. Only one spectrum out of every 16 samples in time was kept. The frequency-dependent component was estimated by the average of the spectrogram over time, the temporal envelope by the average over frequency, and the separable approximation by their outer product.

The analysis of spectrotemporal separability was conducted on 0.4-s segments; this duration was chosen as it corresponds roughly to optimal durations for evoking CMR in humans. The analysis was also repeated for 1-s segments, without significant differences. The similarity between the spectrogram and the separable approximation was quantified using their mutual information. The expected value of the mutual information is positive even when the model is exact, because the carrier is noisy. Therefore, the expected value of the mutual information when the model is exact was computed using Monte-Carlo simulations and subtracted to form the NSI. The full range of NSI in the data set was between -0.3 and 0.7 bits; typical values were between -0.1 and 0.4.

Envelope fluctuations were quantified by the coefficient of variation of the temporal envelope. Because this measure is always positive, it was compared with its expected value for gaussian processes with the same average power spectrum but without temporal modulation, using Monte-Carlo simulations. The FI is their difference, expressed as a dimensionless *t*-value. The full range of FI in the data set was between -50 and 200; typical values were between -5 and 20.

**Physiological recordings.** Recordings were made in primary auditory cortex of five halothane-anaesthetized cats. Anaesthesia was induced by ketamine and xylazine and maintained with halothane (0.25–1.5%) using standard protocols authorized by the committee for animal care and ethics of the Hebrew University—Hadassah Medical School. Stimuli were generated digitally. Noise bands and tones were mixed after digital-to-analog conversion. Signals were presented to the animal through sealed, calibrated earphones (Sokolich). Frequency response was within  $\pm 10$  dB of its mean between 0.1–40 kHz; 0 dB attenuation is 110 dB SPL for tones and 60 dB Hz<sup>-1/2</sup> for noise bands in the range of frequencies used here. To determine CMR, noise bands of varying width (between BF/32 and BF), which were centred arithmetically at BF, were used. Tones at BF were presented at various levels (from below to above masked threshold) in random order. For response to modulated noise, the noise band was multiplied by a 10-Hz sinusoidal or trapezoidal envelope. In all examples shown in this study, except for the top row in Fig. 4, the spectrum level of all the maskers was identical, causing a 3-dB difference in total energy between the modulated and unmodulated maskers. Single neurons were recorded using a glass-coated tungsten electrode and an online spike sorter (MSD, Alpha-Omega). All the neurons were well separated. To determine tone thresholds in noise, the variance-normalized squared differences between the noise and tone plus noise PSTHs were summed over the stimulus duration for each tone level; thresholds are levels above which these sum of squares increased significantly by a  $\chi^2$  test. This approach mimics to some extent the psychophysical paradigm of 2I2AFC in that we compare directly a 'noise interval', characterized by the responses to noise alone, with the 'tone plus noise' interval, characterized by the responses to the combined stimulus.

Received 5 August; accepted 5 November 1998.

1. Pelleg-Toiba, R. & Wollberg, Z. Discrimination of communication calls in the squirrel monkey: "Call detectors" or "cell ensembles"? *J. Basic Clin. Physiol. Pharmacol.* **2**, 257–272 (1991).
2. Steinschneider, M., Arezzo, J. C. & Vaughan, H. G. Jr Tonal features of speech-evoked activity in primate auditory cortex. *Brain Res.* **519**, 158–168 (1990).
3. Wang, X., Merzenich, M. M., Beitel, R. & Schreiner, C. E. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J. Neurophysiol.* **74**, 2685–2706 (1995).

- Suga, N. Philosophy and stimulus design for neuroethology of complex-sound processing. *Phil. Trans. R. Soc. Lond. B* **336**, 423–428 (1992).
- Aertsen, A. M. H. J., Smolders, J. W. T. & Johannesma, P. I. M. Neural representation of the acoustic biotope: on the existence of stimulus-event relations for sensory neurons. *Biol. Cybern.* **32**, 175–185 (1979).
- Hall, J. W., Haggard, M. P. & Fernandes, M. A. Detection in noise by spectro-temporal pattern analysis. *J. Acoust. Soc. Am.* **76**, 50–56 (1984).
- Schooneveldt, G. P. & Moore, B. C. Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration. *J. Acoust. Soc. Am.* **85**, 273–281 (1989).
- Ruderman, D. L. & Bialek, W. Statistics of natural images: scaling in the woods. *Phys. Rev. Lett.* **73**, 814–817 (1994).
- Richards, D. G. & Wiley, R. H. Reverberations and amplitude fluctuations in the propagation of sound in a forest: implication for animal communication. *Am. Nat.* **115**, 381–399 (1980).
- Klump, G. M. & Langemann, U. Comodulation masking release in a songbird. *Hearing Res.* **87**, 157–164 (1995).
- Rhode, W. S. & Greenwood, D. D in *Abstracts of the 18th Association for Research in Otolaryngology Meeting* 127 (St Petersburg Beach, Florida, 1995).
- Schreiner, C. E. & Urbas, J. V. Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hearing Res.* **32**, 49–63 (1988).
- Rauschecker, J. P., Tian, B. & Hauser, M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* **268**, 111–114 (1995).
- Eggermont, J. J. Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hearing Res.* **74**, 51–66 (1994).
- Carlyon, R. P., Buus, S. & Florentine, M. Comodulation masking release for three types of modulator as a function of modulation rate. *Hearing Res.* **42**, 37–45 (1989).

**Acknowledgements.** This work was supported by a grant administered by the Israel Science Foundation. We thank E. Vaadia, M. Abeles, E. Young and A. Aertsen for critical comments to this manuscript.

Correspondence and requests for materials should be addressed to I.N. (israel@music.md.huji.ac.il).

## Kainate receptors mediate synaptic transmission between cones and ‘Off’ bipolar cells in a mammalian retina

Steven H. DeVries\* & Eric A. Schwartz†

\* Department of Ophthalmology and Visual Science, University of Texas Houston Health Science Center, Houston, Texas 77030, USA

† Department of Pharmacological and Physiological Sciences, University of Chicago, Chicago, Illinois 60637, USA

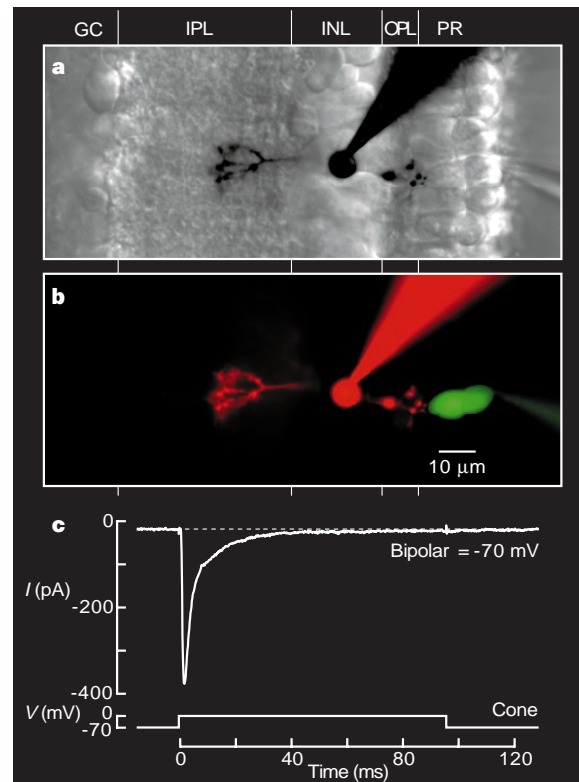
Light produces a graded hyperpolarization in retinal photoreceptors<sup>1,2</sup> that decreases their release of synaptic neurotransmitter<sup>3,4</sup>. Cone photoreceptors use glutamate<sup>5,6</sup> as a neurotransmitter with which to communicate with two types of bipolar cell. Activation of metabotropic glutamate receptors in ‘On’ bipolar cells<sup>7,8</sup> initiates a second-messenger cascade that can amplify small synaptic inputs from cones. In contrast, it is not known how the ionotropic glutamate receptors that are activated in ‘Off’ bipolar cells<sup>9,10</sup> are optimized for transmitting small, graded signals. Here we show, by recording from a cone and a synaptically connected ‘Off’ bipolar cell in slices of retina from the ground squirrel, that transmission is mediated by glutamate receptors of the kainate-preferring subtype. In the dark, a cone releases sufficient neurotransmitter to desensitize most postsynaptic kainate receptors. The small postsynaptic current that persists (<5% of maximum) is quickly modulated by changes in presynaptic voltage. Since recovery from desensitization is slow (the decay time constant is roughly 500 milliseconds), little recovery can occur during the brief (roughly 100-millisecond) hyperpolarization that is produced in cones by a flash of light. By limiting the postsynaptic current, receptor desensitization prevents saturation of the ‘Off’ bipolar cell’s voltage response and allows the synapse to operate over the cone’s entire physiological voltage range.

We studied communication between cones and ‘Off’ bipolar cells in slices of the ground squirrel retina. ‘On’ and ‘Off’ bipolar cells were identified by both electrophysiological and anatomical criteria:

the two types of bipolar cell produce postsynaptic currents of opposite polarity and have axon terminals that end in different halves of the retina’s inner plexiform layer (IPL)<sup>11</sup>. In the experiment illustrated in Fig. 1, the bipolar cell produced an inward postsynaptic current (Fig. 1c) and dye filled an axon terminal in the outer half (or sublamina a) of the IPL (Fig. 1a, b). Thus, both electrophysiology and anatomy showed that the recording was from an ‘Off’ bipolar cell. We studied 99 ‘Off’ bipolar cells which could be subdivided into at least three groups (types b2, b3 and b7 in ref. 12) on the basis of the pattern of their axon terminals in sublamina a. Nonetheless, the electrophysiological properties reported here were essentially the same in all cell pairs. We also observed bipolar cells that produced an outward current and had axon terminals in sublamina b. These were ‘On’ bipolar cells and are not discussed further.

We were immediately struck by the shape of the postsynaptic current (Fig. 1c). A large transient (average  $-228$  pA; range  $-53$  to  $-926$  pA;  $n = 99$ ) quickly declined (decay constant ( $\tau$ ) =  $4.9 \pm 2.9$  ms) to a steady level ( $-9.5 \pm 7.7$  pA or  $4.4 \pm 3.9\%$  of the maximum). We wanted to know how the transient and steady components might contribute to the reliable transmission of small, graded signals produced by a light-stimulated cone photoreceptor.

Our first results showed that  $Ca^{2+}$ -dependent transmitter release was activated throughout a cone’s physiological operating range ( $-60$  to  $-35$  mV) and was responsible for both components of the postsynaptic current. The membrane potential of a cone was stepped from  $-80$  mV to a series of depolarized voltages, first in the presence of  $Ca^{2+}$  and then in a saline solution containing 1 mM  $Cd^{2+}$  (and no  $Ca^{2+}$ ). Exposure to  $Cd^{2+}$  blocked the presynaptic  $Ca^{2+}$



**Figure 1** Synaptic transmission from cone photoreceptors to ‘Off’ bipolar cells. **a**, Sulphorhodamine 101 fluorescence (dark areas) in the bipolar cell is superimposed onto a Nomarski micrograph of the retinal slice. **b**, Sulphorhodamine 101 fluorescence (red) in the bipolar cell and BODIPY 492/515 fluorescence (green) in the cone are superimposed. **c**, Postsynaptic response produced when the cone’s membrane voltage was stepped from  $-70$  to  $0$  mV. Bipolar cells were held at  $-70$  mV. PR, photoreceptors; OPL, outer plexiform layer; INL, inner nuclear layer; GC, ganglion cells.