Article

Phototrophy by antenna-containing rhodopsin pumps in aquatic environments

Ariel Chazan1,12, Ishita Das2,12, Takayoshi Fujiwara3,12, Shunya Murakoshi4,12, Andrey Rozenberg1,12, Ana Molina-Márquez5, Fumiya K. Sano4, Tatsuki Tanaka4, Patricia Gómez-Villegas5, Shirley Larom1, Alina Pushkarev1,10, Partha Malakar6, Masumi Hasegawa3,11, Yuya Tsukamoto3, Tomohiro Ishizuka7, Masae Konno7, Takashi Nagata7, Yosuke Mizuno8, Kota Katayama8,9,Rei Abe-Yoshizumi8, Sanford Ruhman6, Keiichi Inoue7, Hideki Kandori8,9, Rosa León5,Wataru Shihoya4 ✉, Susumu Yoshizawa3 ✉, Mordechai Sheves2 ✉, Osamu Nureki4 ✉ &Oded Béjà1 ✉

Aff1 Faculty of Biology Technion-Israel Institute of Technology Haifa Israel

Aff2 Department of Molecular Chemistry and Materials Science Weizmann Institute of Science Rehovot Israel

Aff3 Atmosphere and Ocean Research Institute The University of Tokyo Chiba Japan

Aff4 Department of Biological Sciences, Graduate School of Science The University of Tokyo Tokyo Japan

Aff5 Laboratory of Biochemistry and Molecular Biology, Faculty of Experimental Sciences, Marine International Campus of Excellence (CEIMAR) University of Huelva Huelva Spain

Aff6 Institute of Chemistry The Hebrew University of Jerusalem Jerusalem Israel

Aff7 The Institute for Solid State Physics The University of Tokyo Chiba Japan

Aff8 Department of Life Science and Applied Chemistry Nagoya Institute of Technology Nagoya Japan

Aff9 OptoBioTechnology Research Center Nagoya Institute of Technology Nagoya Japan

Aff10 Institute for Biology, Experimental Biophysics Humboldt-Universität zu Berlin Berlin Germany

Aff11 Institute for Extra-cutting-edge Science and Technology Avant-garde Research (X-star) Japan Agency for Marine-Earth Science and Technology (JAMSTEC) Kanagawa Japan

Received: 24 August 2022 / Accepted: 31 January 2023

Abstract

Energy transfer from light-harvesting ketocarotenoids to the light-driven proton pump xanthorhodopsins has been previously demonstrated in two unique cases: an extreme halophilic bacterium[1] and a terrestrial cyanobacterium[2]. Attempts to find carotenoids that bind and transfer energy to abundant rhodopsin proton pumps[3] from marine photoheterotrophs have thus far failed[6]. He[4,5,7] detected light energy transfer from the widespread hydroxylated carotenoids zeaxanthin and lutein tothe retinal moiety of xanthorhodopsins and proteorhodopsins using functional metagenomics combined with chromophore extraction from the environment. The light-harvesting carotenoids transfer up to 42% of the harvested energy in the violet-light or blue-light range to the green-light absorbing retinal chromophore. Our data suggest that these antennas may have a significant effect on rhodopsin phototrophy in the world's lakes, seas and oceans. However, the functional implications of ourfindings are yet to be discovered.

Editor's Summary

Light energy transfer from hydroxylated carotenoids to the retinal moiety of light-driven proton pumps is detected.

These authors contributed equally: Ariel Chazan, Ishita Das, Takayoshi Fujiwara, Shunya Murakoshi, Andrey Rozenberg

Main

Proteorhodopsins (PRs) are light-driven proton pumps that were first discovered in abundant marine *Gammaproteobacteria* via the use of metagenomics[7 8]. Since their discovery, numerous homologous proteins have been identified in diverse bacterial and some archaeal and eukaryotic groups from marine and freshwater environments[9 10], and it has been estimated that more than 50% of prokaryotes in the photic zone of the ocean possess these rhodopsins[3 11]. Furthermore, it was suggested that PRs are a major light-harvesting mechanism in the surface ocean[12], enabling PR-bearing microorganisms to lead a photoheterotrophic lifestyle[13 14]

Whereas ma AQ1 AQ2 AQ3 AQ4 AQ5 AQ6 AQ7 AQ8 Rs absorb light preferentially in the green region (green-absorbing PRs (GPRs); absorption maximum (λ) of approximately 525 nm), others show blue-shifted absorption (blue-absorbing PRs (BPRs); λ_{max} of approximately 490 nm)[8 15]. This division is determined mainly by the residue at position 105 (Leu and Met in GPRs eBAC31A08 and MED134, respectively, and Gln in BPRs PalE6 and HOT75m4)[15 16] and enables PRs to absorb light according to water depth, as blue light penetrates deeper in clear oceanic waters[8 16]

The first representative of xanthorhodopsins (XRs), a family of light-driven proton pumps related to PRs, was discovered in the genome of the extreme halophilic bacterium *Salinibacter ruber*[1]. Biophysical characterization revealed that XRs uses two chromophores to absorb light: an all- *rans* retinal moiety covalently bound within the retinal-binding pocket inside the protein, and a 4-ketocarotenoid antenna (salinixanthin) that binds to the protein from the outside and transfers light energy directly to the retinal molecule through a fenestration[17] (that is, an opening in the protein exposing the retinal-binding pocket to the outer environment). This way, salinixanthin enables XRs to absorb a broader range of wavelengths (486 nm and 521 nm in addition to 560 nm absorbed by XRs without the antenna). Another member of the family, GRs from the thylakoid-less terrestrial cyanobacterium *Gloeobacter violaceus*[2], also recruits a 4-ketocarotenoid antenna (echinenone) to transfer light energy to the retinal chromophore within the rhodopsin[18]. The 4-keto group of salinixanthin and echinenone has been shown to be crucial for both binding and energy transfer[17]

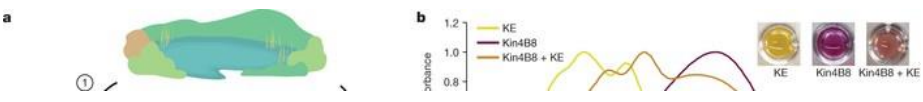
The ability to use a carotenoid antenna requires the presence of the above-mentioned fenestration in the rhodopsin apoprotein that enables the energy transfer between the rings of the two chromophores, and this in turn is facilitated by a conserved Gly residue in transmembrane domain 5 (TM5; Gly156 in *S. ruber* XR)[2]. Environmental XRs (as well as PRs) are variable at this position, with the majority of proteins having either bulky Trp or Phe residues, which render the use of a carotenoid antenna impossible, or the small Gly residue. Despite the fact that natural carotenoids are abundant and diverse, besides the two above-mentioned XRs and a XR from a mountain lake bacterium[19], which binds to the rare xanthophyll nostoxanthin, no other native microbial rhodopsins have been shown to absorb light with predefined carotenoid antennas[4 5] AQ9 Inspired by the known XR–ketocarotenoid light-harvesting systems, we designed a strategy to search for rhodopsin–carotenoid complexes originating from aquatic microorganisms. AQ10 For the first exploratory series of experiments, we picked a rhodopsin with fenestration discovered using functional metagenomics[20 21 22 reshwater lake (Lake Kinneret, Israel) and exposed it to a concentrated chromophore extract from the same lake (Fig. 1a ; see Methods). As the test protein, we chose rhodopsin Kin4B8 (GenBank: OP056329), a member of the XR family with a Gly at XR position 156 from an uncultured *Bdellovibrionota* bacterium. We incubated the purified protein with the Kinneret chromophore extract (KE; Extended Data Fig. 1) and observed a shift in λ_{max} after purification of the complex (Fig. 1b), suggesting a strong binding of specific chromophores. A pronounced circular dichroism spectrum characterized by carotenoid chromophores and probably retinal bands was observed upon addition of the KE cocktail to purified Kin4B8 protein (Fig. 1c), further supporting specific binding of chromophores to the rhodopsin. High-performance liquid chromatography with diode array detector (HPLC-DAD) analysis of the enriched complex showed that the chromophores mainly consisted of the xanthophylls lutein and zeaxanthin (Fig. 1d). AQ11 This result was unexpected as both pigments have a hydroxyl group in position 3 of their β-rings (and also ε-rings in the case of lutein), whereas, initially, only 4-ketocarotenoids were shown to bind to rhodopsins (compare structures in Fig. 1e).

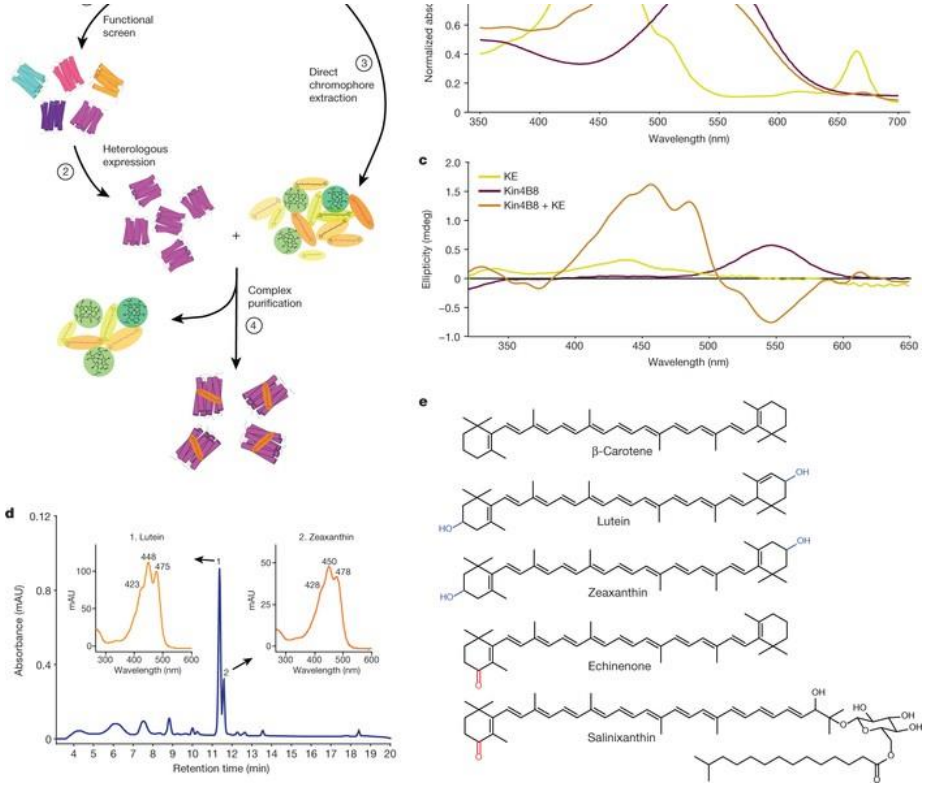
Fig. 1

Environmental xanthophylls bind to a freshwater XR.

, Working pipeline for searching rhodopsin–carotenoid complexes. AQ12 AQ13 (1) Detection of diverse rhodopsins via functional metagenomics. (2) Heterologous expression of selected rhodopsin in *E. coli* and (3) incubation with the chromophore extract derived from the same environment. (4) Rhodopsin is then purified and examined for changes in absorbance spectra. **b**, Absorbance spectra for the KE, the rhodopsin Kin4B8 and the rhodopsin Kin4B8 enriched with KE chromophores. Pictures of the samples are presented in the upper right corner. AQ14 **c**, Circular dichroism spectra with and without KE. **d**, HPLC-DAD chromatogram, registered at 450 nm, of the KE-derived carotenoids bound to the Kin4B8 protein, and the UV–Vis spectra (inset) of the identified peaks. **e**, 2D structures of β-carotene and various hydroxycarotenoids and ketocarotenoids. AQ15

Source data





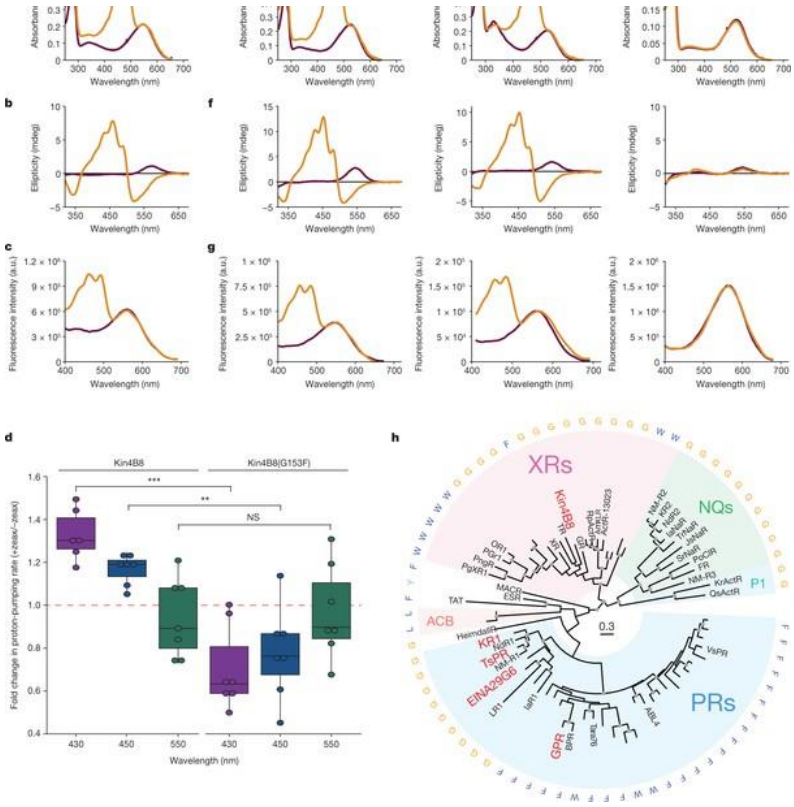
After confirming that commercially available zeaxanthin and lutein also showed binding to Kin4B8 (Fig. 2a,b and Extended Data Fig. 2a,b), we switched to the pure chromophores for further experiments. Fluorescence excitation spectrum (monitored at the retinal emission of 720 nm) of the Kin4B8–xanthophyll complexes exhibited three characteristic peaks of the carotenoid in addition to the expected 560-nm band of the retinal chromophore (Fig. 2c and Extended Data Fig. 2c). The occurrence of the carotenoid bands at the retinal emission indicates that excitation energy from light absorption by the carotenoid is transferred to the excited singlet state of retinal (S_1) and contributes to emission from this state. This observation demonstrates direct energy transfer from zeaxanthin and lutein to the retinal molecule in Kin4B8. On the basis of the fluorescence excitation spectral measurements, we estimate that up to 40% of the light energy harvested by the carotene antenna is transferred to the retinal chromophore (Extended Data Fig. 3a). Excitation energy transfer from carotenoids (as assessed by fluorescence measurements with lutein) was most efficient in the 430–460-nm region (Extended Data Fig. 3b). Further evidence of the energy transfer was obtained by femtosecond transient absorption study (see Methods for details), with the energy transfer efficiency estimated as approximately 55% for the Kin4B8–lutein system (Extended Data Fig. 3c,d). Light-induced difference UV–Vis and Fourier transform infrared spectra of Kin4B8 were similar with and without lutein at 77 K (Extended Data Fig. 4), indicating that carotenoid binding does not affect retinal AQ16 isomerization. Rhodopsin Kin4B8 could also bind to salinixanthin (Extended Data Fig. 5a,b); however, no energy transfer was observed in this case (Extended Data Fig. 5c). Conversely, β-carotene did not show binding to rhodopsin Kin4B8 (Extended Data Fig. 5d,e).

Fig. 2

Biophysical characterization of diverse rhodopsins bound to zeaxanthin.

e, Absorbance change of different rhodopsins upon incubation with (orange) or without (purple) zeaxanthin. OD, optical density. **b f** Circular dichroism spectra with and without lutein. AQ17 **c g**, Fluorescence excitation spectra with (orange) and without (purple) zeaxanthin; emission was monitored at 720 nm. See Extended Data Fig. 2 for the biophysical characterization of the same rhodopsins bound to lutein. **d**, Light-driven initial proton-pumping ratios in Kin4B8 and Kin4B8(G153F) *E. coli* spheroplasts with and without zeaxanthin under violet, blue and green light. The red dashed line represents proton-pumping activity when measured without zeaxanthin. Indicated are significance levels based on *t*-tests with false discovery rate correction: adjusted $***P < 0.001$ and adjusted $**P < 0.01$. NS, not significant. AQ18 AQ19 **h**, Phylogenetic tree of the clade including PRs, XRs and related families. For each one of representative sequences, the residue at the fenestration position is indicated. AQ20 The tree is rooted. Proteins investigated in this study are highlighted in red. ACB, Archaea clade B; ESR, *Exiguobacterium sibiricum* rhodopsin; MACR, marine actinobacterial clade rhodopsin; NQ, NQ chloride and sodium rhodopsin pump; TAT, TAT rhodopsin. See Extended Data Fig. 8a for more details.

Source data



HPLC analysis of retinal oxime produced by the reaction between hydroxylamine and the retinal chromophore in Kin4B8 showed almost equal amounts of all-*trans* and 13-*cis* forms, independently of the light conditions, but the binding of lutein increased the fraction of the all-*trans* form (Extended Data Fig. 6a). Transient absorption changes representing red-shifted (K and O) and blue-shifted (M) photointermediates were observed by laser-flash photolysis of Kin4B8 (Extended Data Fig. 6b). Multi-exponential analysis identified nine photointermediates having different absorption spectra and involving the L and N intermediates with λ_{max} close to the initial state (Extended Data Fig. 6c). A sharp peak at approximately 498 nm was observed when lutein was bound to the protein (Extended Data Fig. 6b, bottom, and Extended Data Fig. 6c), indicating that a large conformational change of the protein occurs during the photocycle that alters the absorption of lutein (Extended Data Fig. 6d). The transient absorption signal was enhanced for Kin4B8 bound to lutein compared with the protein without lutein at excitation wavelengths of less than 500 nm, indicating that energy transfer from lutein to retinal enhances retinal isomerization (Extended Data Fig. 6e).

When measured under violet (410–430 nm) or blue (440–460 nm) illumination, light-dependent outward proton flux in *Escherichia coli* spheroplasts expressing XR Kin4B8 was enhanced upon the addition of zeaxanthin (an addition of approximately 33% in violet light and approximately 17% in blue light; Fig. 2d and Extended Data Fig. 7). The enhancement was not observed under green (550 nm) light illumination. This indicates that violet-light or blue-light energy is absorbed by the xanthophyll antennas and enhances rhodopsin activity. As expected, mutating Kin4B8 Gly153 (XR position Gly156) to Phe blocked the fenestration and abolished carotenoid binding (Extended Data Fig. 5f,g), indicating that the fenestration in rhodopsin Kin4B8 is crucial for xanthophyll binding and energy transfer. In addition, we noticed that blocking the fenestration reduced the pumping activity significantly in the presence of zeaxanthin (Fig. 2d). This decline is probably caused by nonspecific carotenoid adsorption to the surface of spheroplasts, resulting in masking of light. This suggests that the observed increase in proton flux in wild-type Kin4B8 under violet or blue light is in effect underestimated.

Next, we tested whether members of the PR family would also be able to bind to carotenoids. We selected the Gly156-containing PRs EINA29G6 (GenBank: UJI09384) from an uncultured freshwater flavobacterium[22] and KR1 (GenBank: BAN14807) from the marine flavobacterium *Dokdonia eikasta*[23], and GPR 31A08 (GenBank: AAG10475) from a marine gammaproteobacterium[8], which has Phe at XR position 156 instead (Fig. 2h and Extended Data Fig. 8a). Fenestrated PRs EINA29G6 and KR1 indeed demonstrated binding of hydroxylated carotenoids (Fig. 2e,f and Extended Data Fig. 2d,e) and showed direct energy transfer (40% and 35%, respectively) from the xanthophylls to the retinal chromophore (Fig. 2g and Extended Data Figs. 2f and 3a,b). As expected, non-fenestrated GPR 31A08 did not bind to xanthophylls, and hence did not show energy transfer.

Lutein and zeaxanthin are widespread xanthophylls in freshwater and marine environments[24,25,26] (see also HPLC profiles from a freshwater lake and Atlantic coastal waters in Extended Data Fig. 1). Zeaxanthin is readily synthesized by a range of bacteria and is, in particular, one of the dominant carotenoids among *Bacteriodota* (=*Bacteroidetes*)[27], which correlated with the abundance of fenestrated rhodopsins in this group (see Extended Data Fig. 8b). To demonstrate that xanthophylls can function as

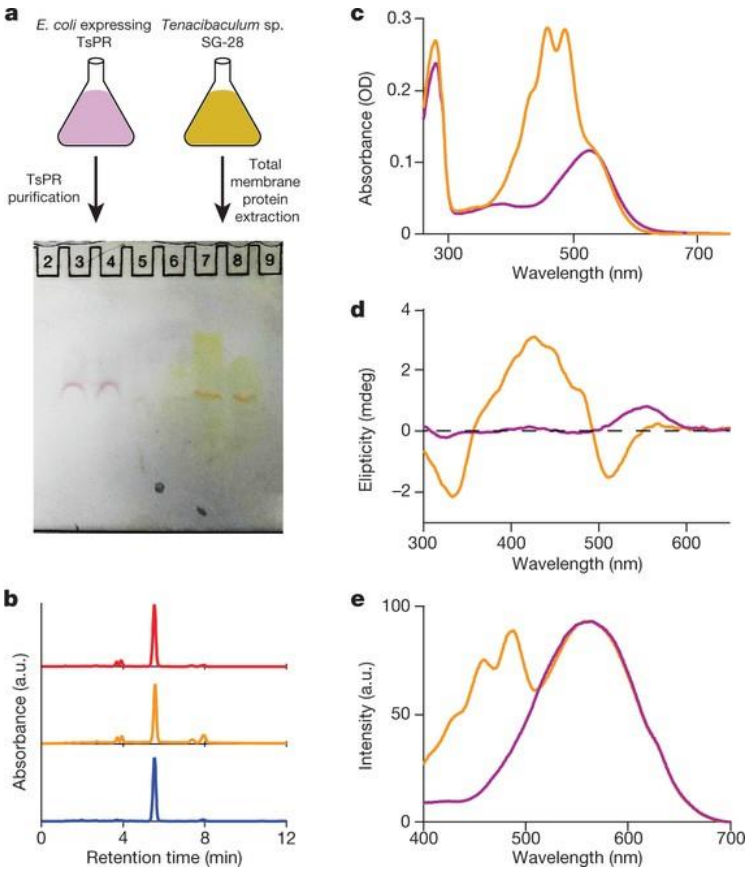
antenna chromophores natively, we turned to a simple cultured system. *Tenacibaculum* sp. SG-28C, a cultured marine flavobacterium, contains a gene encoding only a single rhodopsin (TsPR; GenBank: PQJ23084) in its genome (GenBank: MQVY01000000) and produces zeaxanthin as its sole carotenoid (Fig. 3). We demonstrated that zeaxanthin binds to TsPR in native membranes of *Tenacibaculum* sp. SG-28C (see native gel in Fig. 3a and binding to purified TsPR in Fig. 3c,d) and transfers energy to the retinal molecule (Fig. 3e).

Fig. 3

Characterization of TsPR and carotenoids from *Tenacibaculum* sp. SG-28.

, Gel separation of purified TsPR (expressed in *E. coli*, left) and SG-28 whole-membrane extract (right) on a native PAGE to retain the retinal (in both) and zeaxanthin (in SG-28 only) molecules bound to the rhodopsin protein. For gel source data, see Supplementary Fig. 1 **b**, HPLC profiles of pigments, registered at 475 nm. The red, orange and blue chromatograms are whole carotenoid pigments extracted from SG-28 cells, binding carotenoid pigments to TsPR, and reference samples of pure zeaxanthin, respectively. **c** Absorbance spectra of purified TsPR before (purple) and after (orange) adding whole carotenoid pigments extracted from SG-28 cells. **d**, Circular dichroism spectra with (orange) and without (purple) zeaxanthin. AQ21 , Fluorescence excitation spectra with (orange) and without (purple) zeaxanthin; emission was monitored at 720 nm. AQ22

Source data

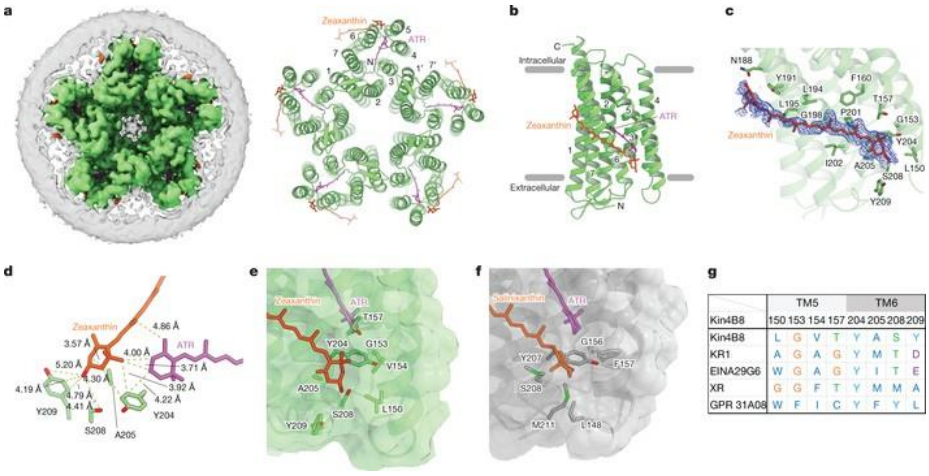


To examine xanthophyll binding to rhodopsins, we determined the cryo-electron microscopy structure of Kin4B8 with zeaxanthin at 2.3 Å resolution (Fig. 4a, Extended Data Fig. 9a–c and Supplementary Tables 1 and 2). Kin4B8 forms a pentamer as other omega rhodopsins[28] (Extended Data Fig. 9d) and the overall structure of Kin4B8 superimposed well on that of the related XR from *S ruber* (Protein Data Bank ID: 3DDL)[29] (Extended Data Fig. 9e). Kin4B8 has conserved structural features observed in other outward proton pumps (Extended Data Fig. 9f). Zeaxanthin lies transversely against the outer surface of TM6 (Fig. 4c and Extended Data Fig. 9c). The hydroxyl ring of zeaxanthin fits into the fenestration in Kin4B8 and forms van der Waals interactions with the retinal β-ionone ring (Fig. 4d,e). The binding mode of zeaxanthin is largely similar to that of salinixanthin in *S. ruber* XRs[29] (Fig. 4f) and would enable energy transfer by a similar mechanism. Of note, the hydroxyl group of zeaxanthin is exposed to the solvent, as the bottom of the fenestration formed by Ser208 and Tyr209 is not closed (Fig. 4e). Homologous residues in KR1 and EinA29G6 are also polar (Fig. 4g), whereas they are hydrophobic in *S. ruber* XRs (Fig. 4f). Together, we hypothesize that the environment at the bottom of the fenestration might be responsible for distinguishing between hydroxylated or ketolated carotenoid selection among various XRs and PRs.

Fig. 4

Structure of the Kin4B8 XR bound to zeaxanthin.

a, Electron microscopy map and pentameric structure of the zeaxanthin-bound Kin4B8, viewed from the extracellular side. **b**, Overall structure of the monomeric unit, with zeaxanthin and the retinal chromophores. ATR, AQ23 XXX. **c**, Zeaxanthin-binding site with the density contoured at 2.0σ. The extended carotenoid is tightly bound to the transmembrane surface of Kin4B8, traversing nearly the entire bilayer, with an inclination of about 60° to the membrane normal. Of the two hydroxyl rings of zeaxanthin, one forms a π–π stacking with Tyr191 and the other binds in a pocket between TM5 and TM6, in close proximity to the β-ionone ring of the retinal. **d** Positional relationship between the hydroxyl ring of zeaxanthin, retinal and surrounding residues. The angle between the chromophore axes is 46°. The angle between the planes of their π-systems is 77°. Of note, the hydroxyl ring forms a π–cation stacking with Ala205, and its dimethyl group is proximal to the retinal. **e f**, Fenestrations in Kin4B8 (**e**) and *S. ruber* XRs (**f**). **g**, Conservation of the residues surrounding the fenestrations in XRs and PRs. AQ24



	TM5				TM6			
Kin4B8	150	153	154	157	204	205	208	209
Kin4B8	L	G	V	T	Y	A	S	Y
KR1	A	G	A	G	Y	M	T	D
EiNA29G6	W	G	A	G	Y	I	T	E
XR	G	G	F	T	Y	M	M	A
GPR 31A08	W	F	I	C	Y	F	Y	L

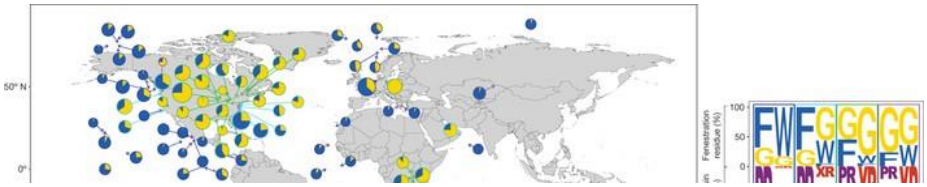
To estimate how widespread the rhodopsin–antenna complexes might be in aquatic environments, we analysed the abundance of different PRs and XRs in available metagenomic datasets. We tentatively assumed that all Gly156-containing PRs and XRs have the potential to bind to antennas, as the structural analysis is not yet conclusive with respect to other common amino acid positions required for carotenoid binding. As the touchstone category, we focused on PRs and XRs with the canonical motif DTE that represent the vast majority of rhodopsins from this clade in the ocean and freshwater, respectively (Fig. 5 and Extended Data Fig. 8c). Among these rhodopsins, on average, 33.1% of PRs and 18.5% of XRs in the ocean and 53.9% of PRs and 81.6% of XRs in the freshwater environments contain the fenestration-enabling Gly156 (Fig. 5). Furthermore, calculations using the *Tara* Oceans marine metagenomic set (Extended Data Fig. 10) show that, on average, 10.3% (±6.90 s.d.) of marine microorganisms per sample in the surface layer contain the fenestration-enabling Gly156 rhodopsins. Prediction of absorption spectra revealed that PR and XR variants with Gly156 show a tendency to absorb in green in the region of relatively longer wavelengths (Extended Data Fig. 8d). At the same time, the fenestration site has not been previously implicated in spectral tuning in rhodopsins and blocking the fenestration in Kin4B8 had no effect on λ_{max} (see above). We thus hypothesize that the observed correlation is an indication that the more red-shifted rhodopsins tend to be fenestrated to allow a broader range of absorbed wavelengths when in complex with carotenoid antennas.

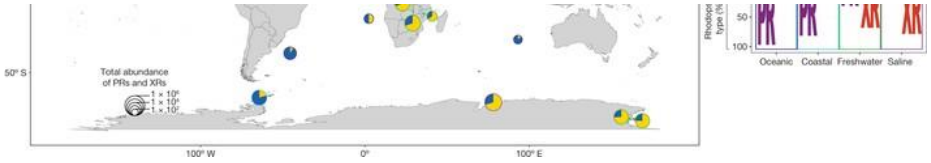
Fig. 5

Global distribution of fenestrated XRs and PRs.

Rhodopsin sequences from JGI metagenomic datasets were classified into PRs and XRs, filtered to include sequences with the DTE motif, and binned into those having bulky residues (Phe or Trp) or Gly at XR position 156. Ratios between fenestrated (yellow) and non-fenestrated (blue) PRs and XRs by location are based on the cumulative abundance for the contigs containing the corresponding rhodopsin genes and are shown as pie charts, with the size of the charts proportional to the total abundance of PR-containing and XR- containing contigs. Locations are subdivided by habitat into: oceanic (blue stroke), coastal (cyan), freshwater (green) and inland saline (purple). The inset shows the ratios between PRs and XRs and the incidence of the different residues at XR position Gly156. See Extended Data Fig. 10 for the distribution of fenestrated marine XRs and PRs based on the *Tara* Oceans data.

Source data





Our study builds on the previous observations that the Gly156-containing XRs from *S. ruber* and *G. violaceus* bind to ketolated carotenoid antennas and the more recent discoveries of a fenestrated actinobacterial XR binding an undefined native carotenoid[30], and a fenestrated proteobacterial XR binding the rare xanthophyll nostoxanthin[19]. Our experiments indicate that this phenomenon is widespread, with the most common xanthophylls being recruited by XRs and even PRs, in freshwater and in the sea. Analysis of the distribution of XRs and PRs in metagenomic datasets further suggests that rhodopsin–carotenoid complexes have a global effect on the amount of sun energy harvested by the microorganisms. Nevertheless, it must be stressed that the exact interactions between the carotenoid and the rhodopsin vary and thus universal structural determinants of carotenoid binding and specificity in fenestrated XRs and PRs are not yet known. It is thus clear that more of these complexes have to be characterized before the hypothesis about their global impact is tested more vigorously.

Methods

Carotenoids

Lutein (PHR1699) and β-carotene (PHR1239) were purchased from Sigma-Aldrich, and zeaxanthin (A132185) from AmBeed. AQ25 Salinixanthin was extracted from purified cell membranes of *S. ruber*-containing XRs as previously described[2] In brief, purified XR was denatured in 2% SDS, followed by lyophilization for 20 h. The lyophilisate was suspended in acetone by 10 min of vortexing followed by centrifugation at 30,000*g*, at 4 °C for 15 min. The resultant residual material mixture after evaporating acetone was run in a silica column, and salinixanthin was eluted and collected with 30/70 (v/v) acetone/hexane mixed solvent after washing several times with 10/90 (v/v) acetone/hexane mixed solvent.

Sample preparation to study carotenoid interaction with rhodopsins

Lutein and zeaxanthin concentrated stock solutions were prepared in DMSO. Approximately equimolar concentration of carotenoids was added to the protein solubilized in 0.05% DDM, 0.3 M NaCl and 0.05 M phosphate buffer at pH 7.5. After incubation of the carotenoids with the proteins for 1 day, the complex was purified using a Ni-resin column to eliminate unbound lutein or zeaxanthin. These samples were used for further spectral studies. For the measurements performed at pH 5.5, 50 mM citrate buffer (with 300 mM NaCl and 0.05% DDM) at pH 5.5 was added to the samples and pH was adjusted by the addition of a small volume of HCl or NaOH if required further.

Reduction of the retinal protonated Schiff-base bond in the Kin4B8–lutein complex

To the purified Kin4B8–lutein complex (in 0.1 M pH 7.5 phosphate buffer with 300 mM NaCl and 0.05% DDM), 0.15 M $NaBH_4$ was added. The reaction was carried out under continuous illumination for 20 min. The light used was filtered through a long pass cut-off filter with $\lambda > 550$ nm (Schott). Then, unreacted $NaBH_4$ was eliminated by dialysing the sample against 0.1 M NaCl and 0.02% DDM. The absorption spectrum of the retinal-protonated Schiff-base-reduced Kin4B8 with lutein is shown in Extended Data Fig. 3c,d

Rhodopsin expression

Kin4B8 XR

Fosmid 4B8 (GenBank: OP056329.1) was selected from our previous metagenomic screen of a fosmid library from lake Kinneret[21]. The 39-kb insert shared high sequence similarity and synteny with metagenomically assembled genomes (MAGs) from undescribed genus UBA2466 (GTDB classification: phylum *Bdellovibrionota*, c__FAC87, o__UBA2466, f__UBA2466, g__UBA2466). The open reading frame of the XR gene, in particular, had a 99.5% and 99.4% identity to assemblies GCA_002359565.1 and GCA_009926805.1, respectively. The gene encoding Kin4B8 XR was cloned into the pBAD plasmid (pBAD TOPO TA Expression Kit, Thermo Fisher Scientific) using forward (5′-AAACCATGGGTTCTGCAACTACACTAACGCTG-3′) and reverse (5′-GCGTTTAAACTTAGTGATGGTGATGATGAGCGGGTAGTGAGCC-3′) primers containing NcoI and MssI (PmeI) restriction sites and a 6× His-tag at the C terminus. The point mutation G153F (Gly156 in *S. ruber* XR), prepared to block the fenestration in Kin4B8, was obtained using the NEB Q5 site-directed protocol (https://nebasechanger.neb.com/) with primers 5′-ATTAATGTGGTTCGTCTTAAGTACCGTGCCCTTCC-3′ and 5′-CGAGCGCCAGCTTGGCTT-3′.

DH10b *E. coli* cells with pBAD-Kin4B8 or pBAD-Kin4B8(G153F) plasmids, were grown in LB supplemented with 50 μg ml^{-1} ampicillin at 220 r.p.m. at 37 °C. When the OD_{600} reached 0.6, expression was induced using a 0.1% final concentration of L-arabinose (A3256, Sigma-Aldrich). The induced culture was grown at 220 r.p.m. overnight (more than 16 h) at 30 °C. Then, the culture was incubated with 20 μM all-*trans* retinal for more than 4 h in the dark.

KR1

The plasmid and expression protocol was adapted from ref. [23]. In brief, BL21 (DE3) *E. coli* cells with the pET-21a-KR1 plasmid were grown on LB with 50 μg ml^{-1} ampicillin at 220 r.p.m. at 37 °C. When the OD_{600} reached 0.5, expression was induced using a 0.5 mM final concentration of isopropyl β-D-1-thiogalactopyranoside (IPTG; 1758-1400, Inalco Pharmaceuticals), and all-*trans* retinal at a final concentration of 20 μM was supplemented. The induced culture was grown at 220 r.p.m. overnight (more than 16 h) at 30 °C.

EINA29G6 PR and GPR

Fosmid EIN29G6 (GenBank: MW650840.1) was isolated from a freshwater sample and the proton-pumping activity for the encoded PR has previously been assessed by us[22]. The 36- kb fosmid is syntenic with scaffolds from freshwater flavobacterial MAGs, in particular, to assembly GCA_016868755.1 (GTDB taxonomy: phylum *Bacteroidota*; c__Bacteroidia; o__Flavobacteriales; f__Schleiferiaceae; g__TMED14) with which it shows 92.6% DNA identity in the PR open reading frame. Single DH10b *E. coli* colonies harbouring pBAD-EINA29G6-PR[22] or pBAD-GPR plasmids[8] were used to inoculate LB media supplemented with 100 μg ml^{-1} ampicillin and 0.001% or 0.2% (EINA29G6 and GPR, respectively) final concentration of L-arabinose. The culture was incubated in a deep 96-well plate overnight (more than 16 h) at 750 r.p.m. at 30 °C. Then, the pooled plate content was incubated with 20 μM all-*trans* retinal for more than 4 h in the dark.

TsPR

Single C41 (DE3) *E. coli* colonies harbouring pET21a-TsPR (TsPR; GenBank: PQJ23084.1)[31] were used to inoculate LB media supplemented with 100 μg ml^{-1} ampicillin. Before protein expression, transformants were cultured at 37 °C in 2× YT medium (NaCl 5 g l^{-1}, Bacto tryptone 16 g l^{-1} and Bacto yeast extract 10 g l^{-1}, pH 7.0) with 100 μg ml^{-1} ampicillin until the absorbance at 660 nm reached 0.5. Protein expression was then induced at 37 °C for 4 h in the dark by adding 0.1 mM IPTG (Sigma-Aldrich) and 10 μM all-*trans*-retinal.

Protein purification

Rhodopsin-expressing *E. coli* cultures were centrifuged at 5,000*g* for 15 min at 4 °C and the pellet was resuspended in a buffer containing 50 mM Tris HCl pH 8.0, 5 mM $MgCl_2$ and 0.1 mM PMSF (P7626, Sigma-Aldrich). The sample was disrupted by using a microfluidizer for 10 passes at 60 psi. Then, the sample was centrifuged at 5,000*g* for 15 min at 4 °C to pellet undisrupted cells or large cell debris. Membranes were collected by centrifuging the sample at 37,000*g* for 1 h at 4 °C and resuspended in a buffer containing 50 mM Tris HCl pH 8.0, 300 mM NaCl, 5 mM imidazole, 5 mM $MgCl_2$, 10% glycerol and 2% DDM final concentration. The sample was incubated for more than 2 h at 4 °C with gentle rotation and a second centrifugation at 37,000*g* for 1 h at 4 °C was performed. The supernatant was then incubated with Ni-Beads (31103, Cube Biotech) for 1 h. Beads were washed on a gravity column using a buffer containing 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 0.05% DDM and 20–50 mM imidazole. Protein was eluted from the column using a buffer containing 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 0.05% DDM and 250 mM imidazole. Eluted protein was washed using Amicon 100-kDa cut-off (UFC910008, Millipore) with storage buffer 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol and 0.05% DDM. The protein was then flash frozen in liquid N_2 and stored at −80 °C.

TsPR-expressing *E. coli* cells were resuspended in 7 ml of buffer containing 50 mM Tris-HCl pH 8.0, and 500 mM NaCl. The cells were disrupted by sonication (Branson SFX 250 Digital Sonifier, Branson Ultrasonics) on ice-cold water for 5 min. Crude membranes were obtained by ultracentrifugation at 4 °C at 106,800*g* for 30 min (Optima XPN-90 Ultracentrifuge with a SW 32Ti rotor, Beckman Coulter) and solubilized with 1% DDM. The solubilized TsPR was collected by ultracentrifugation at 4 °C at 106,800*g* for 30 min and purified by a HisTrap FF Ni^{2+}-NTA affinity chromatography column (GE Healthcare) at room temperature (approximately 25 °C). The purified sample was concentrated, and its buffer was replaced with a new buffer containing 25 mM MOPS pH 7.2, 500 mM NaCl and 0.1% DDM using an Amicon Ultra Filter 30-kDa cut-off (Millipore) by centrifugation at 4 °C at 5,000*g* for 20 min (MX-305, Tomy Seiko).

Native PAGE

Tenacibaculum sp. SG-28 cells were cultured and collected as described above. The cells were resuspended in 20 ml of buffer containing 50 mM Tris-HCl pH 8.0 and 500 mM NaCl and disrupted in the French press (EmulsiFlex-C3, Avestin) at 8 °C. Unbroken cells were removed by centrifugation (3,300*g* for 5 min; MX-305, Tomy Seiko). Then, crude membranes were obtained by ultracentrifugation at 4 °C at 106,800*g* for 30 min (Optima XPN-90 Ultracentrifuge with a SW 32Ti rotor, Beckman Coulter) and discarding the supernatant. The membrane fraction was solubilized with 1.0% (w/v) DDM (Dojindo Lab.) and concentrated using an Amicon Ultra Filter (30 kDa; Millipore) by centrifugation at 15 °C at 4,000*g* for 60 min (MX-305, Tomy Seiko).

For native PAGE, 15 μl TsPR purified protein (heterologously expressed in *E. coli*; A_{520} of 0.5), or 15 μl cell membrane fraction sample, was mixed with 15 μl of the buffer (Native Sample Buffer, Bio-Rad). Protein separation was performed using 4–15% precast gel (4–15%, 10 wells; Mini-Protean TGX Stain-free Gels, Bio-Rad). Gel electrophoresis was carried out in a Mini-PROTEAN Tetra System (Bio-Rad), using 10× Tris/glycine buffer (Bio-Rad) as a running buffer at 4 °C at 8 mA for 11 h, under dark conditions. An uncropped image of the gel is available in Supplementary Fig. 1

Environmental chromophore sampling and extraction

Lake Kinneret sampling

Of water, 200 l was sampled on 24 August 2021 in Lake Kinneret, Station A (32° 49′ 16.6764″ N, 35° 35′ 20.5512″ E). Water

samples were then concentrated by a tangential flow filtration system (N06-E100-05-S, Repligen) with a 100 kDa cut-off after pre-filtration through a mesh net. Concentrated water samples were disrupted using a microfluidizer (M-110S, Microfluidics Corporation) by applying 30–80 PSI for five cycles. The sample was then freeze-dried using a lyophilizer (Coolsafa 110-4, ScanVac) for approximately 48 h. Chromophore extraction was done directly on the dried material using hexane extraction[32] In brief, dried samples were resuspended in 10 ml of acetone by applying extensive pipetting and vortexing. Hexane and 10% NaCl were added to the mixture in a 2:2:1 ratio (acetone to hexane to 10% NaCl). The mixture was vortexed and then centrifugedat 3,000*g* at 4 °C for 3 min. The hexane (top) layer was then transferred to a separate falcon and the process was repeated till the hexane phase became colourless. Combined hexane fractions were then dried using N_2 gas and reconstituted in 5 ml absolute ethanol.

Atlantic coast sampling

Of seawater, 3 l was sampled on 07 June 2021 on the coast of Huelva, southwest of Spain (37° 12′ 50.6″ N, 6° 59′ 48.3″ W) and filtered through 0.7-µm microfibre filters (GF/F). To extract the chromophores, the filter was incubated with cold acetone (−20 °C) overnight. Then, the sample was centrifuged and the supernatant was filtered through a 0.45-µm filter and dried using N_2 gas. **AQ26**

Tenacibaculum sp. SG-28 chromophore extraction

Tenacibaculum sp. SG-28 cells were cultured in 500 ml half-strength ZoBell's 2216E medium under light conditions at 25 °C. Cultured cells were collected by centrifugation (4,400*g* for 10 min at 20 °C) and then washed twice in 100 mM NaCl. Intracellular pigments were extracted with acetone/methanol (7:3 v/v). Air-dried pigments were resuspended in 30 ml of the buffer containing 25 mM MOPS pH 7.2, 500 mM NaCl and 0.1% DDM, and filtered through a 0.22-µm pore size filter (Advantech) to remove insoluble pigments. The filtered solution was concentrated using an Amicon Ultra Filter 30-kDa cut-off (Millipore) by centrifugation at 15 °C at 4,000*g* for 60 min (MX-305, Tomy Seiko).

Binding of chromophore extract to rhodopsin proteins

Kin4B8 and KE

Of Kin4B8 purified protein, 10 mg was mixed with 5 ml KE at absorbance of approximately 5 and incubated overnight with gentle rotation at 4 °C. Approximately 2.5 ml of Ni-Beads were added to the mixture for 4 h at 4 °C. The protein was then washed extensively using a storage buffer and eluted using the same buffers used for the initial purification.

TsPR and *Tenacibaculum* sp. SG-28 TE

Purified TsPR and TE were mixed in a ratio of 450 µl (A_{520} of 0.5) to 600 µl (A_{450} of 1.5) and incubated overnight with gentle rotation at 4 °C. The mixed samples were purified using a HisTrap FF Ni^{2+}-NTA affinity chromatography column (Takara Bio) at 4 °C. The concentration and replacement of the buffer was carried out in the same way as purification of TsPR.

Chromatographic analysis of carotenoids

N_2-dried lyophilized KE (10 mg) and Atlantic coastal chromophore extract samples were resuspended in 1 ml of methanol, transferred into 2-ml screw-top polystyrene tubes with 0.5 g of 0.5-mm glass beads under N_2 atmosphere, treated for 2 min in a Genie disruptor and incubated overnight at −20 °C. After centrifugation at 9,700*g* for 10 min, the supernatant was analysed in a Hitachi Chromaster HPLC equipped with a DAD detector, using an RP-18 column at a flow rate of 1 ml min^{-1}. Solvents were a mixture of acetonitrile/water (9:1 v/v; solvent A) and pure ethyl acetate (solvent B). The gradient elution program was as follows: 0–16 min at 0–60% for A; 16–30 min at 60% for A; 30–35 min at 100% for A. **AQ27** The column temperature was kept at 25 °C, and the injection volume was 100 µl. Chromaster Hitachi control software was used for data processing. The salinixanthin standard was prepared by purification from *S. ruber*. Other standards were purchased from Sigma-Aldrich or DHI Lab (Hørsholm).

To identify pigments bound to TsPR, pigment analysis of the TE and the TE–TsPR mixture were performed by HPLC. Of TE (A_{450} of 1.0) or TE–TsPR mixture (OD_{520} of 0.5), 100 µl was added to 500 µl of methanol and evaporated at 30 °C. Air-dried samples were resuspended in 500 µl of acetone/methanol (1:9 v/v) and stored at −20 °C for 3 h. These solutions were analysed in a Shimadzu HPLC (model Prominence-i LC-2030) equipped with a PDA detector, using a Kinetex 5 µm EVO C18 100A column (Phenomenex) at a flow rate of 2 ml min^{-1}. Solvent was a mixture of methanol/water (9:1 v/v). The column temperature was kept at 60 °C, and the injection volume was 20 µl. LabSolutions software was used for data processing. Commercial zeaxanthin (Fujifilm Wako Pure Chemical) was used as a standard.

Absorption spectroscopic measurements

UV−Vis absorption spectral measurements of rhodopsins (Kin4B8, Kin4B8(G153F), KR1, EINA29G6 and GPR) with the carotenoids (lutein, zeaxanthin, β-carotene and salinixanthin) and KE were taken with a Cary 8454 UV−vis spectrophotometer (Agilent Technologies). Protein and carotenoid concentrations were kept in the approximately 2–6-µM range for absorption spectral measurements.

UV−Vis absorption spectral measurements of TsPR and the TE–TsPR mixture were taken with a UV-2600 spectrophotometer (Shimadzu).

Circular dichroism spectroscopic measurements

Circular dichroism (CD) spectroscopic measurements of rhodopsins Kin4B8, Kin4B8(G153F), KR1, EINA29G6 and GPR with the carotenoids (lutein, zeaxanthin, β-carotene and salinixanthin) and KE were performed with Chirascan CD spectrometer (Applied Photophysics). A quartz cell of 1-cm path length was used for the measurements. CD spectra were recorded with 2.1-nm bandwidth resolution in a 1-nm interval.

CD spectroscopic measurements of TsPR and TE–TsPR were performed with a J-725 spectrometer (Jasco) at room temperature (approximately 25 °C). A quartz cell of 1-cm path length was used for the measurements. CD spectra were recorded with 1-nm bandwidth resolution in a 1-nm interval.

Low-temperature UV–Vis and FTIR spectroscopic analysis

Low-temperature UV-visible and Fourier transform infrared (FTIR) spectroscopies were performed as previously described[33] except for minor modifications of the reconstitution into the membrane process. In brief, purified Kin4B8 samples, with and without lutein, were reconstituted into a membrane composed of POPE:POPG (3:1 mol/mol) (POPE:POPG; Avanti) with a protein-to-lipid molar ratio of 1:20, by removing DDM with Bio-Beads (SM-2, Bio-Rad). Kin4B8 samples in POPE:POPG liposomes were washed repeatedly with a buffer containing 5 mM NaCl and 2 mM NaH_2PO_4 (pH 7.25) and collected by ultracentrifugation for 20 min at 222,000*g* at 4 °C. The lipid-reconstituted Kin4B8 sample was finally suspended in the same buffer and then placed on a BaF_2 window to prepare dry-layer thin film. Kin4B8 films were hydrated with 1 µl H_2O before measurements. The sample was then placed in the cell of a cryostat (Optistat, Oxford Instruments) mounted in the UV–Vis spectrometer (V-750, Jasco) and a FTIR spectrometer (Carry670, Agilent Technologies) and cooled to 77 K. Each difference spectrum was calculated from two spectra constructed from 128 interferograms with a resolution of 2 cm^{-1}. Illumination with 540-nm light for 30 s through an interference filter (KL-54 interference filter, Toshiba) at 77 K converted Kin4B8 into the K intermediate, which reverted to the original state upon illumination at more than 590-nm light (R-61 cut-off filter, Toshiba) for 30 s, as evidenced by a mirror image of the difference spectra. Averages of 100 and 60 experiments were conducted for the spectra of Kin4B8 with and without lutein, respectively.

HPLC analysis of retinal isomers

The HPLC analysis of retinal isomers was performed as previously described[34] with minor modifications. Purified samples in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1% DDM and 10% glycerol were kept at room temperature in the dark before the experiments. For light adaptation, samples were illuminated with green light (530 ± 5 nm) for 60 s, followed by incubation in the dark for 60 s. A 75-µl sample was mixed with 280 µl of 90% (v/v) methanol aqueous solution and 25 µl of 2 M hydroxylamine (NH_2OH) to convert retinal chromophore into retinal oxime, and then the retinal oxime was extracted with 800 µl of *n*-hexane. A 200 µl of the extract was injected into an HPLC system equipped with a silica column (particle size of 3 µm, 150 × 6.0 mm; Pack SIL, YMC), a pump (PU-4580, Jasco) and a UV–Vis detector (UV-4570, Jasco). The solvent for the mobile phase was -hexane containing 15% ethyl acetate and 0.15% ethanol and the flow rate was 1.0 ml min^{-1}. The molar composition of the retinal isomers in the sample was determined with the molar extinction coefficient at 360 nm for each isomer (all-*trans*-15-*syn*: 54,900 M^{-1} cm^{-1}; all-*trans*-15-*anti*: 51,600 M^{-1} m^{-1}; 13-*cis*-15-*syn*, 49,000 M^{-1} cm^{-1}; 13-*cis*-15-*anti*: 52,100 M^{-1} cm^{-1}; 11-*cis*-15-*syn*: 35,000 M^{-1} cm^{-1}; and 11-*cis*-15-*anti*: 29,600 M^{-1} cm^{-1}).

Laser-flash photolysis

For the laser-flash photolysis spectroscopy, Kin4B8 with and without lutein was solubilized in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1% DDM and 10% glycerol. Optical density of the rhodopsin was adjusted to approximately 0.4–0.5 (protein concentration of approximately 0.2–0.25 mg ml^{-1}) at the absorption maximum wavelengths. The laser-flash photolysis measurement was conducted as previously described. The nanosecond second harmonics of Nd–YAG laser (λ = 532 nm, INDI40, Spectra-Physics) was used for the excitation of Kin4B8. The transient absorption spectra were obtained by monitoring the intensity change of white light from a Xe-arc lamp (L9289-01, Hamamatsu Photonics) passed through the sample with an ICCD linear array detector (C8808-01, Hamamatsu Photonics). To increase the signal-to-noise ratio, 60–100 spectra were averaged, and the singular-value decomposition (SVD) analysis was applied. To measure the time evolution of transient absorption change at specific wavelengths, the output of a Xe-arc lamp (L9289-01, Hamamatsu Photonics) was monochromated by monochromators (S-10, Soma Optics) and the change in the intensity after the photoexcitation was monitored with a photomultiplier tube (R10699, Hamamatsu Photonics). To increase signal-to-noise ratio, 100–200 signals were averaged.

To measure the transient absorption change of pyranine due to proton release and uptake by wild-type Kin4B8, the protein was solubilized in 50 mM sodium-phosphate, 300 mM NaCl and 0.05% DDM at pH 7.5. Nanosecond pulses from an optical parametric oscillator (basiScan, Spectra-Physics) pumped by the third harmonics of Nd–YAG laser (λ = 355 nm; INDI40, Spectra-Physics) were used for the excitation of Kin4B8 at different wavelengths (λ_{exc} = 432, 458, 473, 488, 547 and 595 nm). The pulse energy was lowered to 0.3 mJ cm^{-2} to keep the linearity between the number of the absorbed photon and the transient absorption change.

Fluorescence spectroscopic measurements

Fluorescence emission and excitation spectral measurements were taken on a Jobin Yvon-Spex Fluorolog-3 spectrofluorometer, which is composed of a 450W Xe-lamp as the light source, double-grating monochromator in the excitation and emission positions, and a photomultiplier tube detector (R928P). The slit width of both the emission and the excitation channels were mostly kept at 8 or 10 nm. Absorbance of the samples used for fluorescence measurements was kept within OD of 0.3 with respect to the carotenoid absorption maximum. However, collected spectral profiles were further corrected for the internal absorption effect. Retinal fluorescence emission intensity of Kin4B8 was good enough for measurements at pH 7.5, but for KR1, GPR and

EINA29G6, it was too weak to be detectable at this pH. Therefore, for these latter three rhodopsins, all the fluorescence measurements were done at pH 5.5 (below the pK_a of the retinal PSB counter-ions) to get detectable fluorescence signals. Excitation spectra were sampled at 720 or 730 nm to avoid the strong Raman bands that mask the retinal fluorescence. Fluorescence excitation spectra were scaled to the respective retinal absorption spectral for the calculation of excitation energy transfer efficiency of the respective donor–acceptor pairs. A representative figure (for the Kin4B8–lutein system) showing the absorption and fluorescence excitation spectral profiles scaled together is presented in Supplementary Fig. 2. The following equation, relating the fluorescence excitation spectral profile to the corresponding absorption spectral profile, was used for estimation of quantum efficiency of excitation energy transfer from the carotenoid antenna to the retinal chromophore: Exc(λ) $(1 - 10^{A}) \times (A + \phi A)/A$, where Exc($\lambda$) and A ($= A + A$) are the excitation spectrum and absorbance of the complex, respectively; A_c and A_r are the absorption spectrum of bound carotenoid, and the retinal component, respectively; and ϕ is the quantum efficiency of energy transfer[35]

Fluorescence emission and excitation spectra of TsPR and TE–TsPR were measured on the RF-6000 spectrofluorometer (Shimadzu) at room temperature (approximately 25 °C). The slit width of both the emission and the excitation channel was 5 nm. Scattered excitation light was blocked by a long-pass filter (more than 510 nm; Asahi Spectra). Absorbance of the samples used for fluorescence measurements for the TsPR sample was 0.5 (A525), and for the TE–TsPR sample was 0.6 (A485). The fluorescence measurements were done at pH 5.5 and the excitation spectra were sampled at 720 nm. The following equation relating the fluorescence excitation spectral profile to the corresponding absorption spectral profile was used to correct the inner filter effect: $0^{Ex/2}$, where F_{ideal} is the ideal fluorescence signal spectrum absence of the inner filter effect, and F_{obs} and A_{Ex} are the excitation and absorbance spectra, respectively[36]

Femtosecond transient absorption measurements of Kin4B8–lutein

Femtosecond transient absorption measurements were performed using a hybrid Ti:Sapphire laser system operating at 1,000 Hz. A description of the laser system, pump and probe pulse generation and its detection have been detailed elsewhere[37] Femtosecond transient absorption measurements were performed on an optical cell with a 0.5-mm quartz window and 0.4-mm path length. The sample cell was rotated rapidly to prevent photodegradation. In addition, sample integrity was monitored by measuring its absorbance throughout the measurement.

Treatment of the native rhodopsin with $NaBH_4$ yielded a reduced retinal protonated Schiff-base bond (RPSB), which in turn abolishes energy transfer from the carotenoid, yet maintaining a similar retinal-binding site[38,39]. In the case of the Kin4B8(reduced RPSB bond)–lutein complex, the S_2 state lifetime of lutein was 140 fs (τ_D), which was shortened to 60 fs (τ_{DA}) for the Kin4B8–lutein complex, confirming the energy transfer process. The efficiency of energy transfer (ϕ_{ET}) is calculated by using $\phi_{ET} = 1 - \tau_{DA}/\tau_D$. Refer to Extended Data Fig. 3a,b for the femtosecond decay kinetics of the S_2 state of lutein in the complex.

Proton-pumping measurements

The protocol was adapted from ref. [15] with minor modifications. In brief, 200 ml of rhodopsin-expressing *E. coli* cells was centrifuged at 3,600*g* for 10 min and resuspended into 20 ml of 30 mM Tris-HCl pH 8.0 and 20% sucrose. Of lysozyme (L6876, Sigma-Aldrich), 200 µg was added to the cell suspension and gently rotated for 1 h at room temperature. Spheroplasts were centrifuged at 3,600*g* for 15 min at room temperature and the pellet was resuspended with 2 ml of 100 mM KPi pH 7.0, 20 mM $MgSO_4$ and 20% sucrose supplemented with 4 mg of DNase I (DN25, Sigma-Aldrich). The solution was injected using a syringe into 50 ml of 50 mM KPi pH 7.0 for 15 min at 37 °C. Na-EDTA was added to the mixture to a final concentration of 10 mM and the solution was stirred for another 15 min. Then, $MgSO_4$ was added at a final concentration of 15 mM, followed by another 15 min of stirring. Cell debris was removed by centrifugation at 3,600*g* for 10 min and vesicles were collected by centrifugation at 16,000*g* for 1 h. Spheroplasts were resuspended with 2 ml of 100 mM KPi pH 7.0 and 10 mM $MgSO_4$. Addition of 20 µM zeaxanthin to approximately 1 ml (half of the volume) followed by overnight incubation at 4 °C with gentle rotation. Spheroplasts were then washed twice with 100 mM KPi pH 7.0 and 10 mM $MgSO_4$, and three times with unbuffered solution (10 mM NaCl, 10 mM $MgSO_4 \cdot 7H_2O$ and 100 µM $CaCl_2$). Samples were kept in the dark until the pH was stabilized and illuminated using Leica 1177 equipped with a 150W halogen lamp and filtered with violet, blue and green interference filters (420–430, 445–455 and 560–600 nm, respectively); light intensity was approximately 225, 335 and 990 µmol m^{-2} $^{-1}$, measured after the interference violet, blue and green filters, respectively (measured using LI-COR Biosciences LI-250A light meter). pH was monitored using a LAQUA F-72G pH/ION meter (HORIBA scientific) equipped with a 9618S-10D pH microelectrode. pH results were converted into proton concentration $[H^+]$ 10^{p}, and the fold change in the proton-pumping rate over the first 10 s of illumination was calculated using the following equation:

$$\text{Fold change} = \frac{([H^+]_{10} - [H^+]_0)_+}{([H^+]_{10} - [H^+]_0)_{hi}}$$

Where the fold change in the proton-pumping rate is represented by the delta in proton concentration over the first 10 s of illumination of the result with zeaxanthin divided by the results without zeaxanthin[40]. Differences in pumping activity between the wild-type variant and the fenestration mutant were assessed with tailed *t*-test after exploring normality with the Shapiro–Wilk test The *P* values were false discovery rate adjusted.

Protein expression and purification for structural analysis

pBAD-Kin4B8 was transfected in *E. coli* C41 (Rosetta). The transformant was grown in LB supplemented with 50 µg ml^{-1} ampicillin and 10 µg ml^{-1} at 220 r.p.m. at 37 °C. When the OD_{600} reached 0.6, expression was induced using a 0.1% final concentration of L-arabinose (A3256, Sigma-Aldrich). The induced culture was grown at 120 r.p.m. overnight (more than 16 h) at 25 °C. Then, the pooled plate content was incubated with 20 µM all-*trans* retinal for more than 4 h in the dark. The collected cells were disrupted by sonication in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl and 10% glycerol. The crude membrane fraction was collected by ultracentrifugation at 180,000*g* for 1 h. The membrane fraction was solubilized in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% DDM, 10% glycerol and 10 mM imidazole for 2 h at 4 °C. The supernatant was separated from the insoluble material by ultracentrifugation at 180,000*g* for 20 min and incubated with Ni-NTA resin (Qiagen) for 30 min. The resin was washed with 10 column volumes of wash buffer containing 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.03% DDM, 10% glycerol and 25 mM imidazole. The resin was incubated with 1 column volume of wash buffer containing 100 µM lutein or zeaxanthin. Then, the resin was washed with 5 column volumes of wash buffer. The protein was eluted in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.03% DDM, 10% glycerol and 300 mM imidazole. The eluate was dialysed against buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.03% DDM). The protein was concentrated to 40 mg ml^{-1} using a centrifugal filter device (Amicon 50 kDa MW cut-off), and frozen until crystallization.

X-ray crystallographic analysis of the Kin4B8–lutein complex

The protein bound to lutein was reconstituted into monoolein at a weight ratio of 1:1.5 (protein to lipid). The protein-laden mesophase was dispensed into 96-well glass plates in 30-nl drops and overlaid with 800 nl precipitant solution, using a Gryphon robot (ARI), as previously described[41]. Crystals of Kin4B8 were grown at 20 °C in precipitant conditions containing 25% PEG550MME, 100 mM HEPES-NaOH pH 7.0 and 180 mM potassium thiocyanate. The crystals were harvested directly from the LCP using micromeshes (MiTeGen) and frozen in liquid nitrogen, without adding any extra cryoprotectant. **AQ33**

X-ray diffraction data were collected at the SPring-8 beamline BL32XU with an EIGER X 9M detector (Dectris), using a wavelength of 1.0 Å. In total, 282 small-wedge (10° per crystal) datasets using a 15 × 10-µm^2 beam. The collected images were processed using KAMO[42] with XDS[43], and 45 datasets were indexed with the consistent unit cell parameters. After correlation coefficient-based clustering using normalized structure factors followed by merging using XSCALE[44] with outlier rejections implemented in KAMO, the second largest cluster consisting of 106 datasets was selected for the downstream analyses, because it gave the highest inner-shell and outer-shell CC1/2. The Kin4B8 structure was determined by molecular replacement with PHASER[44], using the model calculated by AlphaFold2 (ref. [45]). Subsequently, the model was rebuilt and refined using Coot[46] and phenix.refine[47]. The asymmetric unit contains two protomers (molA and molB). The final model of Kin4B8 contained residues 6–213 and 216–259 (for molA), 3–259 (for molB), retinal, 14 monoolein molecules and 25 water molecules. The lutein molecule was not modelled.

Cryo-EM single-particle analysis of the Kin4B8–zeaxanthin complex

For the cryo-EM grid preparation of the zeaxanthin-bound Kin4B8, the protein was reconstituted in nanodiscs. Kin4B8, MSP2N2 and SoyPC were mixed at a molar ratio of 1:4:200, respectively, and incubated on ice for 30 min. Detergents were removed by adding Bio-Beads SM2 (Bio-Rad) to 40 mg ml^{-1}, followed by gentle agitation. The Bio-Beads were replaced with fresh ones every 2 h, twice in total. The third batch of Bio-Beads (equal amount) was incubated at 4 °C overnight. The Bio-Beads were then removed by passage through a PolyPrep column (Bio-Rad), and the lysate was ultracentrifuged before size-exclusion chromatography. Ultracentrifuged samples were purified by size-exclusion chromatography on a Superose 6 Increase 10/300 GL column (GE Healthcare), equilibrated with buffer containing 20 mM Tris-HCl pH 8.0 and 150 mM NaCl. The peak fractions of the protein were collected and concentrated to an absorbance (A_{280}) of 12.5, using a centrifugal filter unit (50 kDa molecular weight cut-off; Merck Millipore). Of protein, 3 µl was loaded onto glow-discharged holey carbon grids (Quantifoil Au 300 mesh R1.2/1.3), after which, these were plunge-frozen in liquid ethane, using a Vitrobot Mark IV (Thermo Fischer Scientific).

Cryo-EM imaging was collected on a Titan Krios at 300 kV, using a Gatan K3 Summit detector. Images were obtained at a dose rate of about 8.0 e Å^{-2} s^{-1}, with a defocus ranging from −0.6 to −1.6 µm. Total exposure time was 8 s, with 40 frames recorded per micrograph. A total of 9,256 videos were collected. Raw micrographs in super-resolution mode were binned by 2× and corrected for both full-frame and local motion with the patch motion correction in cryoSPARC[48 49]. Contrast transfer function (CTF) parameters were estimated with patch CTF estimation in cryoSPARC. Particles were initially picked from a small fraction with Gaussian blob picking and subjected to 2D classification. Class averages showing reasonable features of the Kin4B8 pentamer in various orientations were selected as templates for template-based particle picking. Particles from these class averages generated an ab initio model in cryoSPARC. For each full dataset, extracted particles were downsampled to 3.32 Å, followed by two rounds of 2D classification to remove 'junk' particles. 3D classification into two classes with the ab initio model as a reference was performed using C5 symmetry. After multiple rounds of 3D classification, particles were re-extracted with the bin-1 pixel size of 0.83 Å and a box size of 360 pixels. Multiple rounds of local CTF refinement and non-uniform refinement[50] were performed using cryoSPARC. Finally, the 429,741 particles in the best class were reconstructed using non-uniform refinement, resulting in a 2.33 Å resolution reconstruction, with the gold-standard Fourier shell correlation (FSC = 0.143) criteria[51]. The processing strategy is described in Extended Data Fig. 9a . The quality of the map was sufficient to build a model manually in Coot[46 52]. Model building was performed based on the crystal structure of Kin4B8. We manually readjusted the model into the density map using Coot and refined it using phenix.real_space_refine (v.1.19)[47 53]. Finally, we refined the model using servalcat[54]. The final model of Kin4B8 contained residues 3–256, retinal, zeaxanthin and 24 water molecules. Detailed parameters are listed in Supplementary Table 2

Bioinformatic analyses

Rhodopsin sequence analysis

Rhodopsin sequences were classified using a curated set of reference sequences from the superclade uniting PRs, XRs, NQ rhododopsins and related clades (PR–XR–NQ superclade). The reference sequences were used (1) to build a custom HMM profile (the PR–XR–NQ profile) that was used to collect rhodopsin sequences from environmental datasets, and (2) to create a database against which query sequences were searched with usearch_global from usearch (v.11.0.667)[55] for fine-grained classification. The residues at active positions were determined based on the alignment against the PR–XR–NQ profile. Absorption maxima were predicted for the query sequences using the BLASSO machine learning model published previously[56] with a custom wrapper workflow available at https://github.com/BejaLab/BLASSO-Rhodopsin

Metagenomic datasets

Three metagenomic datasets were recruited to quantify the distribution of fenestrated and non-fenestrated rhodopsins: Ocean Microbial Reference Catalog (OM-RGC) v.2 (ref. [57]), metagenomes distributed via JGI Integrated Microbial Genomes & Microbiomes (IMG/M)[58] and Genomes from Earth's Microbiomes (GEM) catalogue of MAGs[59]. Abundance values for individual rhodopsin sequences were taken as the per-gene coverage as supplied in OM-RGC v.2 and as per-contig coverage as supplied in IMG/M. In addition, for OM-RGC data, abundances per cell were estimated by dividing rhodopsin abundances by the averaged abundance of the ten single-copy markers[60] as identified in the catalogue by orthology groups COG0012, COG0016, COG0018, COG0172, COG0215, COG0495, COG0525, COG0533, COG0541 and COG0552. For IMG/M, only studies with unrestricted data utilization status were recruited. Individual metagenomic studies were binned into four habitat types: open ocean, coastal marine (estuaries, lagoons, channels and near-shore habitats), freshwater and inland saline water bodies. For the map visualization of the ratios between fenestrated and non-fenestrated PRs and XRs, abundances were summed for individual samples per location with a resolution of 1°. Maps were created with ggplot2 (ref. [61]), scatterpie (https://cran.r-project.org/web/packages/scatterpie/) and other R packages.

Phylogenetic analysis

The phylogenetic tree was constructed for a reference set of sequences composed mostly of characterized sequences by aligning them with mafft v.7.505 (ref. [62]) in automatic mode, trimming the alignment with trimAl v.1.4.1 (ref. [63]) and building the phylogeny with iqtree2 v.2.2.0.3 with ultrafast bootstrap[64 65]

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-05774-6

Extended data

is available for this paper at https://doi.org/10.1038/s41586-023-05774-6

Supplementary information

The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-05774-6

Acknowledgements

The paper is dedicated to the memory of Dieter Oesterhelt (1940–2022). AQ34 We thank J. K. Lanyi for commenting on the manuscript; M. Shalev-Benami for help in optimizing protein purification; G. Tzuri and T. Isaacson for sharing materials; J. Anton for providing *S. ruber* for the isolation of salinixanthin standard; E. Martinez-Montes, director of the 'Marismas del Odiel' Natural Park, for support in obtaining the Atlantic Sea water sample; the Yigal Allon Kinneret Limnological Laboratory (KLL) for technical help with Lake Kinneret sampling; the Interuniversity Institute for Marine Sciences (IUI) in Eilat for providing access to their pier as well as workspace for the primary processing of the Red Sea samples; and the Israel Nature and Parks Authority for the permit to sample in Ein Afek reserve. This work was supported by the Israel Science Foundation (grant 3592/19 to O.B.), the Institute for Fermentation Osaka (W.S.), JSPS KAKENHI (grants 18H04136 and 22H00557 to S.Y., JP21H01875 and JP20K21383 to K.I., 19H05777 to W.S., 21H04969 to H.K. and 21H05037 to O.N.), MEXT Advancement of Technologies for Utilizing Big Data of Marine Life (grant JPMXD1521474594 to S.Y.), MEXT KAKENHI, Grant-in-Aid for Transformative Research Areas (B) 'low-energy manipulation' (grant JP20H05758 to K.I.), the Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research) from the Japan Agency for Medical Research and Development (AMED) under grant number JP19am0101070 (support number 1627 to W.S.), Agencia Estatal de Investigación/FEDER, UE (grant number 2019-110438RB-C22 to R.M.L.), and Binational Science Foundation (grants 2016102 and

2020105 to S.R.). S.R. holds the Lester Aronberg Chair in Chemistry. M.S. holds the Katzir-Makineni Chair in Chemistry. O.B. holds the Louis and Lyra Richmond Chair in Life Sciences. AQ35

Author contributions

A.C. and A.P. conceived the project, performed environmental sampling and functional metagenomics. A.C. performed carotene extraction, protein biochemistry, carotene binding and light-dependent proton pumping. A.R. performed bioinformatics. S.L. performed molecular biology. I.D. and M.S. performed absorption, emission and CD spectroscopies. T.F., M.H., Y.T. and S.Y. performed absorption, emission and carotene characterization of the PR-containing flavobacterial isolate. S.M., F.K.S., T.T., W.S. and O.N. performed structural analysis. A.M.-M., P.G.-V. and R.M.L. performed carotene characterization from environmental samples and from rhodopsin-bound carotenes. P.M. and S.R. performed ultrafast spectroscopy. T.I., M.K., T.N. and K.I. performed laser-flash photolysis. Y.M., K.K., R.A.-Y. and H.K. performed low-temperature UV–Vis and FTIR spectroscopy. O.B. coordinated the project. A.C. and O.B. wrote the paper with input from all authors.

Peer review

Peer review information *Nature* thanks Valentin Gordeliy and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Data availability

All data are available in the main text or the Supplementary Information. The sequence of fosmid Kin4B8 was deposited in GenBank under accession number OP056329. Atomic coordinates of the crystal structure of Kin4B8 have been deposited in the Protein Data Bank under 7YTB. The density map and structure coordinate of the cryo-EM structure of the zeaxanthin-bound Kin4B8 have been deposited in the Electron Microscopy Data Bank and the Protein Data Bank with accession numbers EMD-35143 and 8I2Z, respectively. Source data are provided with this paper.

Code availability

The code used for the bioinformatic analyses is available from the GitHub repository (https://github.com/BejaLab/antenna) and the data are deposited in the Figshare repository (https://doi.org/10.6084/m9.figshare.20502384).

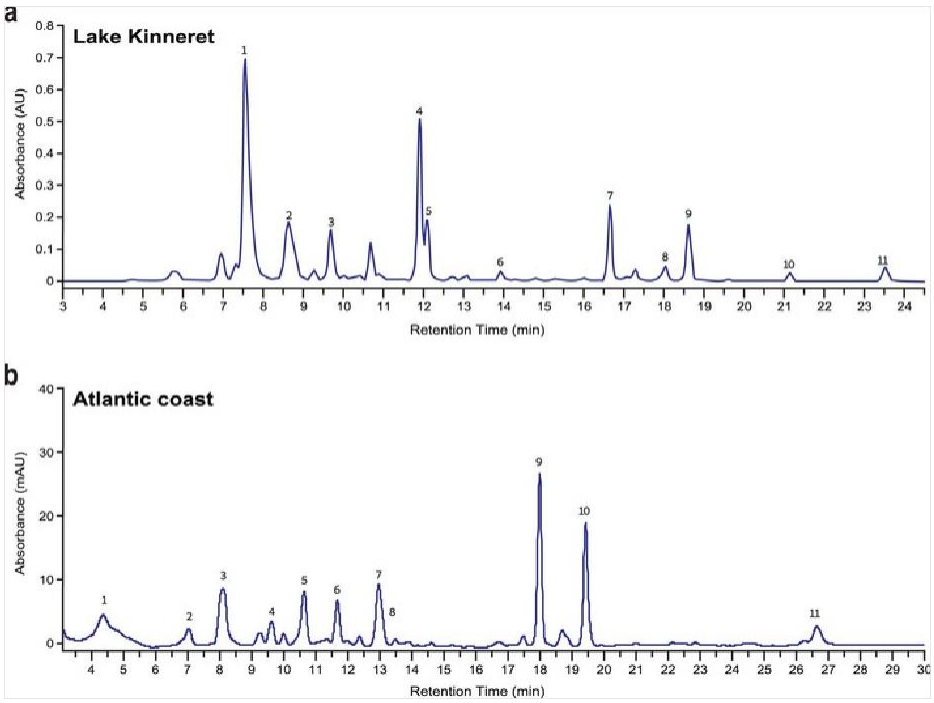
Competing interests The authors declare no competing interest.

Extended data figures and tables

Extended Data Fig. 1

Characterization of Lake Kinneret and Atlantic coast chromophore extracts.

a, HPLC profile of Lake Kinneret chromophore extract. Main peaks correspond to myxoxanthophyll (1), dinoxanthin (2), diatoxanthin (3), lutein (4), zeaxanthin (5), canthaxanthin (6), chlorophyll *b* (7), chlorophyll *a* (8), echinenone (9), pheophytin *a* (10), and β-carotene (11). **b**, HPLC profile of Atlantic coast chromophore extract. Main peaks correspond to chlorophyll *c* (1), siphonaxanthin (2), fucoxanthin (3), violaxanthin (4), diadinoxanthin (5), diatoxanthin (6), lutein (7), zeaxanthin (8), chlorophyll *b* (9), chlorophyll *a* (10), and β-carotene (11). Chromophores registered at 450 nm.

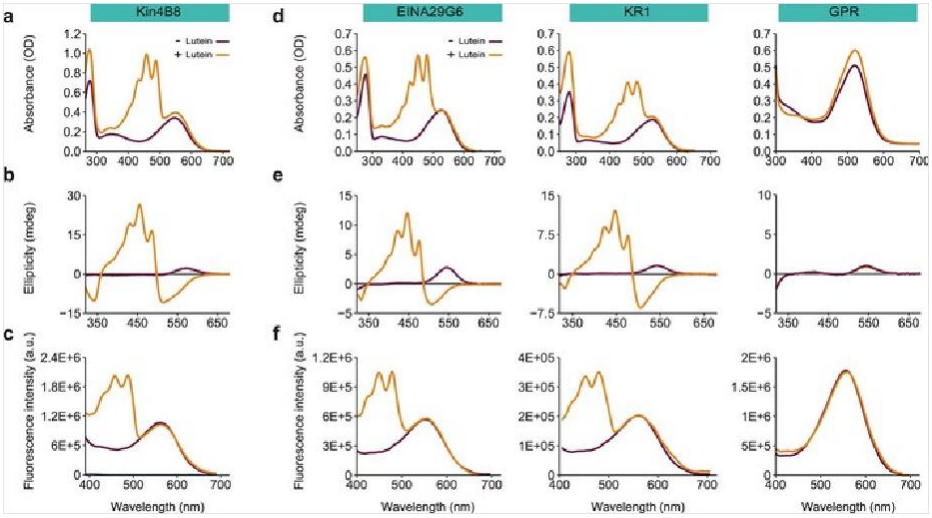


Extended Data Fig.

Spectroscopic characterization of diverse rhodopsins bound to lutein.

a and **d**, Absorbance change of different rhodopsins upon incubation with lutein. **b** **e**, CD spectra with and without lutein. **c** **f** Fluorescence excitation spectra with and without lutein; emission monitored at 720 nm.

Source data



Extended Data Fig. 3

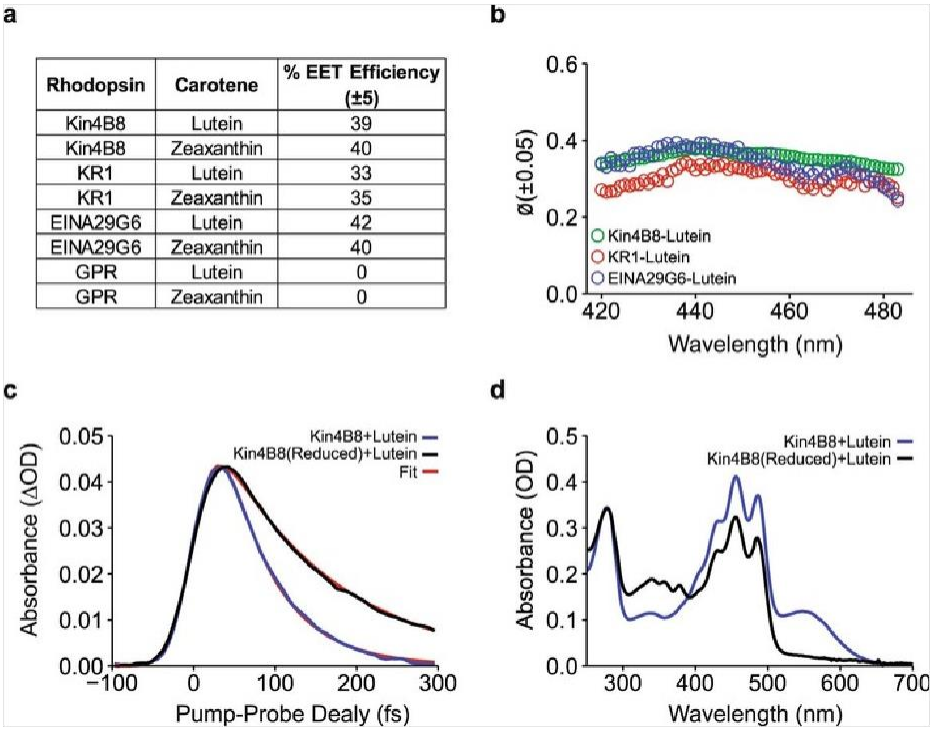
Quantum efficiency of excitation energy transfer (EET) from xanthophylls to different rhodopsins and ultrafast spectroscopy characterization of lutein in complex with Kin4B8.

, Calculated quantum efficiency of EET from lutein and zeaxanthin to different rhodopsins in complex. **b**, Quantum efficiency of EET from lutein to different rhodopsins in complex, as a function of wavelength. , S_2 state decay of lutein. Blue — Kin4B8-lutein, black — Kin4B8 (reduced RPSB bond)-lutein, and red- their fit. X-axis presents the delay between pump and probe. Y-axis shows probe absorption difference in the presence and absence of pump pulse. Kinetic data were fitted with a function convolving 40 fs gaussian IRF and a mono-exponential decay. **d**, Reduction of retinal protonated Schiff-base by $NaBH_4$ in Kin4B8-lutein complex. Blue — absorption spectrum of Kin4B8-lutein before reduction. Black — absorption spectrum of Kin4B8-lutein after reduction.

Source data

a

Rhodopsin	Carotene	% EET Efficiency (±5)
Kin4B8	Lutein	39
Kin4B8	Zeaxanthin	40
KR1	Lutein	33
KR1	Zeaxanthin	35
EINA29G6	Lutein	42
EINA29G6	Zeaxanthin	40
GPR	Lutein	0
GPR	Zeaxanthin	0

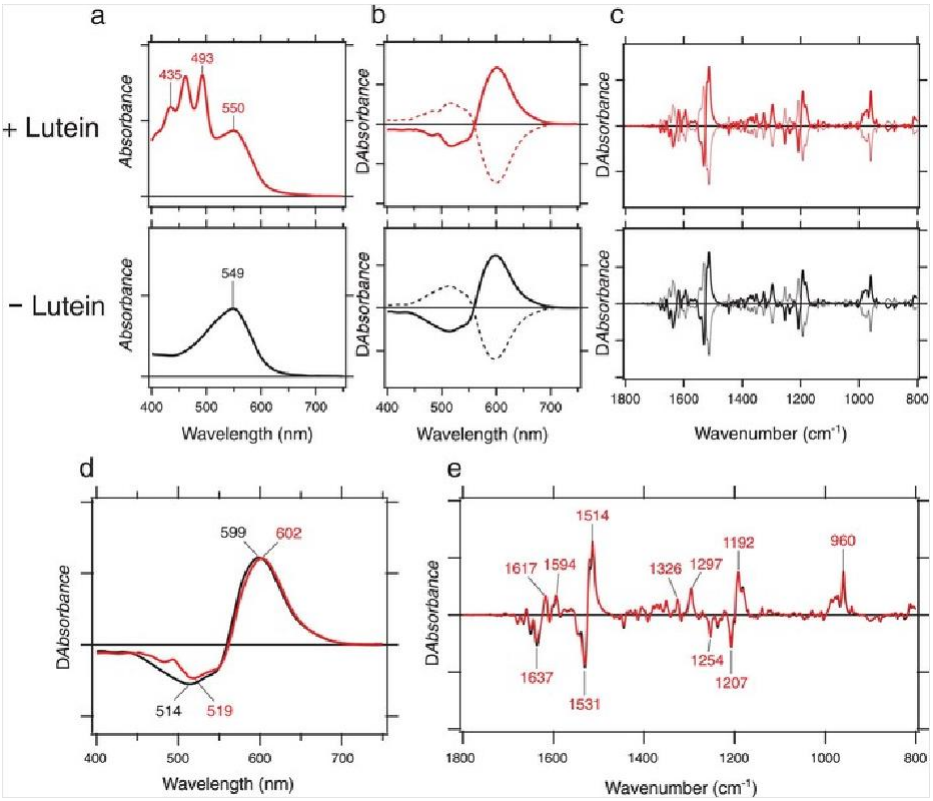


Extended Data Fig. 4

FTIR Influence of lutein on the retinal photoisomerization in Kin4B8 at 77 K.

a, UV-visible absorption spectra of lipid-reconstituted Kin4B8 with (top) and without (bottom) lutein at 77 K. One division of the y-axis corresponds to 0.5 absorbance units. **b**, Difference UV-visible spectra upon illumination of Kin4B8 with (top) and without (bottom) lutein. Hydrated films of lipid-reconstituted Kin4B8 were first illuminated at 540 nm light (solid lines), followed by illumination at >590 nm light (broken lines) at 77 K. Solid and broken lines are mirror-imaged, indicating photochromic properties for Kin4B8 and the K intermediate. One division of the y-axis corresponds to 0.05 absorbance units. **c**, Difference FTIR spectra upon illumination of Kin4B8 with (top) and without (bottom) lutein. Hydrated films of lipid-reconstituted Kin4B8 with H_2O were first illuminated at 540 nm light (solid lines), followed by illumination at >590 nm light (dotted lines) at 77 K. One division of the y-axis corresponds to 0.002 absorbance units. **d** **e**, Light-induced difference UV-visible (**d**) and FTIR (**e**) spectra of Kin4B8 with (red) and without (black) lutein, where positive and negative signals originate from the K intermediate and unphotolyzed Kin4B8, respectively.

Source data

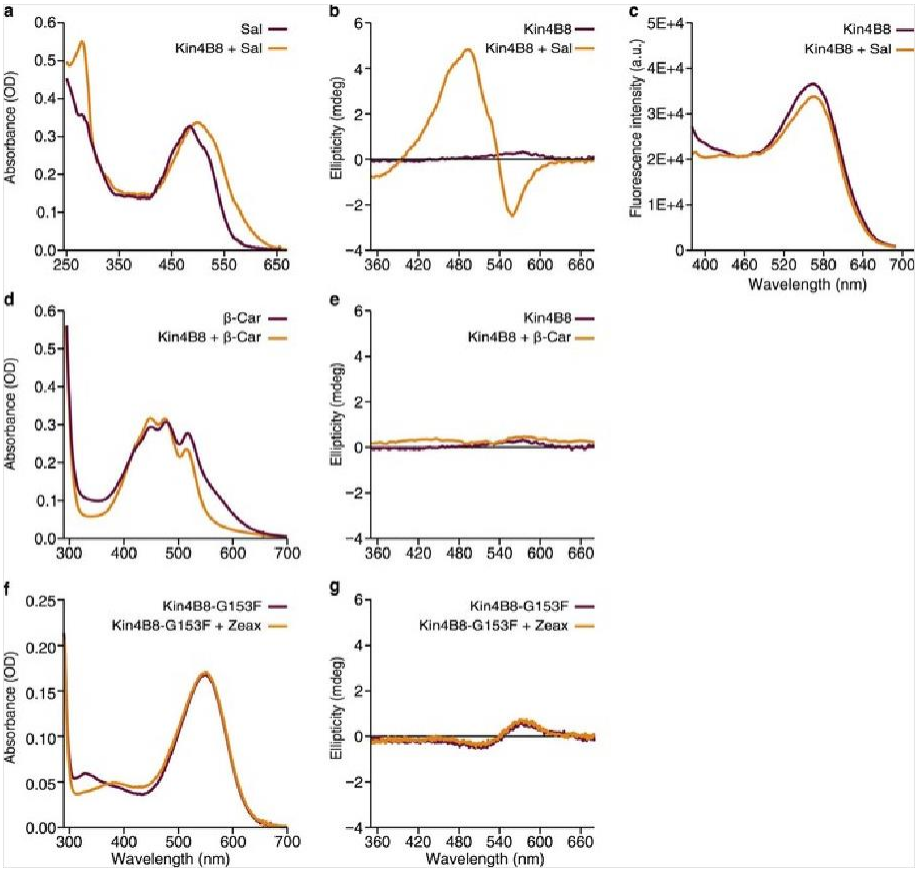


Extended Data Fig. 5

Spectroscopic characterization of Kin4B8 and Kin4B8-G153F bound to carotenoids.

, Absorption spectra of Kin4B8 with salinixanthin (Sal). **b**, CD spectra of Kin4B8 with and without salinixanthin. , Fluorescence excitation spectra of Kin4B8 with and without salinixanthin; emission monitored at 720 nm. **d**, Absorption spectra of Kin4B8 with β-carotene (β-car). **e**, CD spectra of Kin4B8 with and without β-carotene.**f**, .Absorption spectra of Kin4B8-G153F with zeaxanthin (Zeax). **e**, CD spectra of Kin4B8-G153F with and without zeaxanthin.

Source data

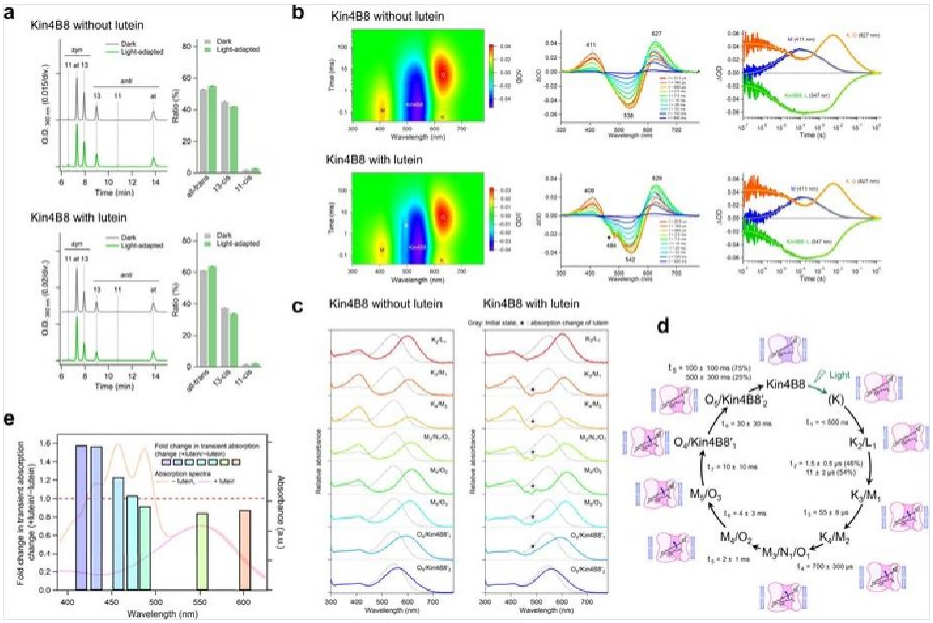


Extended Data Fig. 6

The photocycle of Kin4B8.

a, Chromatogram of HPLC analyses (left) and the compositions of the retinal isomers (right) in Kin4B8 with (top) and without (bottom) lutein under the dark (gray) and light-adapted (green) conditions. at, 11, 13, syn, and anti indicate all-*trans*, 11-*cis*, 13-*cis syn*, and *anti* configurations, respectively. **b**, Two-dimensional plot of transient absorption change (left), transient absorption spectra at different time points (middle), and time course of the transient absorption change (right) of Kin4B8 without (top) and with (bottom) lutein. Peaks derived from the absorption change of lutein are indicated by an asterisk. **c**, Absorption spectra of the photointermediates of Kin4B8 with (left) and without (right) lutein. Peaks derived from the absorption change of lutein are indicated by an asterisk. **d** Photocycle model of Kin4B8. Conformational change of rhodopsin affects the structure of lutein (blue arrow) from K_3/M_1 to $O_4/Kin4B8'_2$ **e**, The ratios of transient absorption change in Kin4B8 with and without lutein at different excitation wavelengths (415, 432, 457, 473, 487, 552, and 601 nm) (bars colored according to the color of excitation light). The absorption spectra of Kin4B8 without (pink line) and with (orange line) lutein were overlaid. The red dashed line indicates no difference between with and without lutein.

Source data

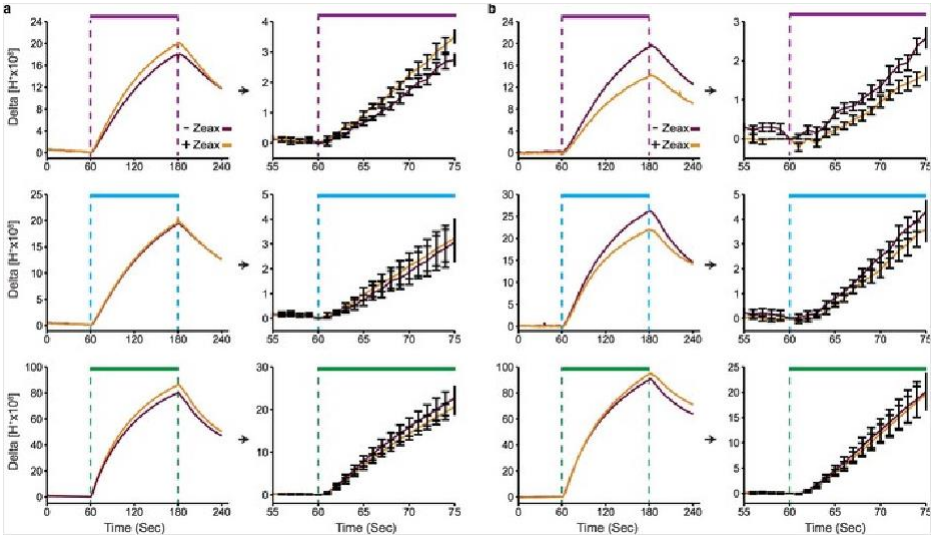


Extended Data Fig. 7

Light-induced proton-pumping activity of Kin4B8 and Kin4B8-G153F.

a and **b**, Monitoring of pH changes in *E. coli* spheroplasts suspension expressing Kin4B8 or Kin4B8-G153F, respectively, with and without zeaxanthin. The spheroplasts were illuminated with violet (430 nm), blue (450 nm), or green (550 nm) light for 2 min (indicated by the colored bars). An enlarged plot of the first 15 s of illumination is displayed to the right of each measurement. The presented traces are the average of six or more independent biological replicates (Error bars represent the SEM).

Source data



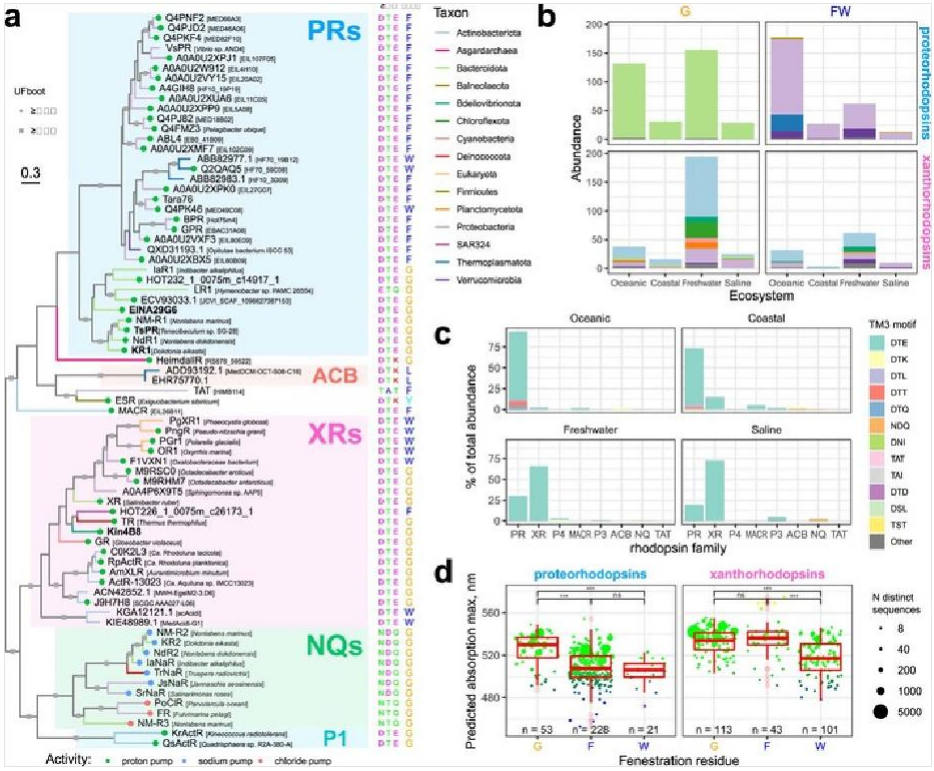
Extended Data Fig. 8

Diversity and distribution of PRs and XRs with (G) and without (FW) fenestration among different prokaryotic phylaacross four environments.

, Maximum likelihood phylogenetic analysis of the PR-XR-NQ clade based on representative protein sequences. Characterized ion pumps are indicated with dots, terminal branches are colored by the corresponding phylum. Major clades with more than one representative are highlighted and labeled. The tree is outgroup rooted. **b**, Distribution of PRs and XRs with the canonical TM3 motif DTE among genomes assigned to different taxa, with (G) and without (FW) fenestration. The analysis is based on GEM genomes and the numbers are summarized per operational taxonomic unit (OTU). The colors are as in panel (a). **c**, Relative abundance of different families of the clade across four habitats based on the metagenomic data from IMG/M. Only families with a total relative abundance of >0.1% are shown. **d**, Predicted absorption maxima for PRs and XRs with the three most frequent residues at the fenestration position. Individual observation corresponds to an average absorption maximum predicted with the rhodopsin BLASSO model for sequences with the same 24 residues of the retinal binding pocket[56]. The sequences from OM-RGC, IMG/M and GEM were pooled together.

The size of the dots is proportional to the number of distinct rhodopsin domain sequences and the color approximates the predicted mean absorption spectra. Statistical differences between the groups were assessed with Dunn's test with FDR correction. Significance levels are indicated with asterisks: *** – adjusted *p*-values < 0.001. Abbreviation of family names in (A) and (C): ACB – Archaea clade B, ESR – *Exiguobacterium sibiricum* rhodopsin, NQ – NQ sodium and chloride pumps, MACR – marine actinobacteria clade rhodopsins, PR – proteorhodopsins, TAT – TAT rhodopsins, XR – xanthorhodopsins, P1 – unnamed clade including QsActR, KrActR and related rhodopsins, P3 and P4 – currently unnamed clades.

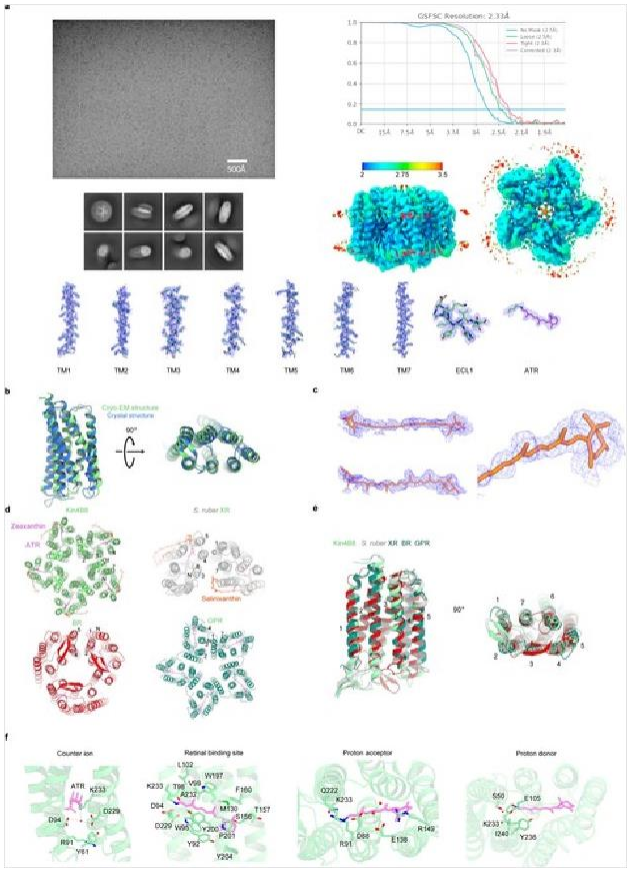
Source data



Extended Data Fig. 9

Structural features of Kin4B8.

a, Cryo-EM single-particle analysis of the zeaxanthin-bound Kin4B8. **b**, Comparison of the cryo-EM and crystal structures of Kin4B8. **c**, Density of the zeaxanthin, contoured at 2.0σ, which allows unambiguous identification of the molecule. In particular, resolution of the half of zeaxanthin proximal to the fenestration is high enough to permit identification of the dimethyl group of its hydroxyl ring. **d** Comparison of the oligomeric structures of Kin4B8, *S. ruber* XR (PDB ID: 3DDL), BR (PDB ID: 1C3W), GPR (PDB ID: 7B03). With the ECL1 sheet inside, Kin4B8 forms a hexamer with aligned directions to the membrane in the crystal packing. The pentameric structure would reflect a physiological condition, in contrast to the previously reported head-to-tail dimer of *S. ruber* XR. **e**, Structural comparison of Kin4B8 with *S. ruber* XR (PDB ID: 3DDL), bacteriorhodopsin (BR) (PDB ID: 1C3W), and GPR (PDB ID: 7B03), with root mean square deviations (RMSD) of 1.44, 1.96, and 2.61 Å, respectively. Notably, the N-terminal region (residues 6–11) and ECL1 form a 3-stranded antiparallel β-sheet, as in *S. ruber* XR and other omega rhodopsins. **f**, Key rhodopsin proton pump motifs in Kin4B8.

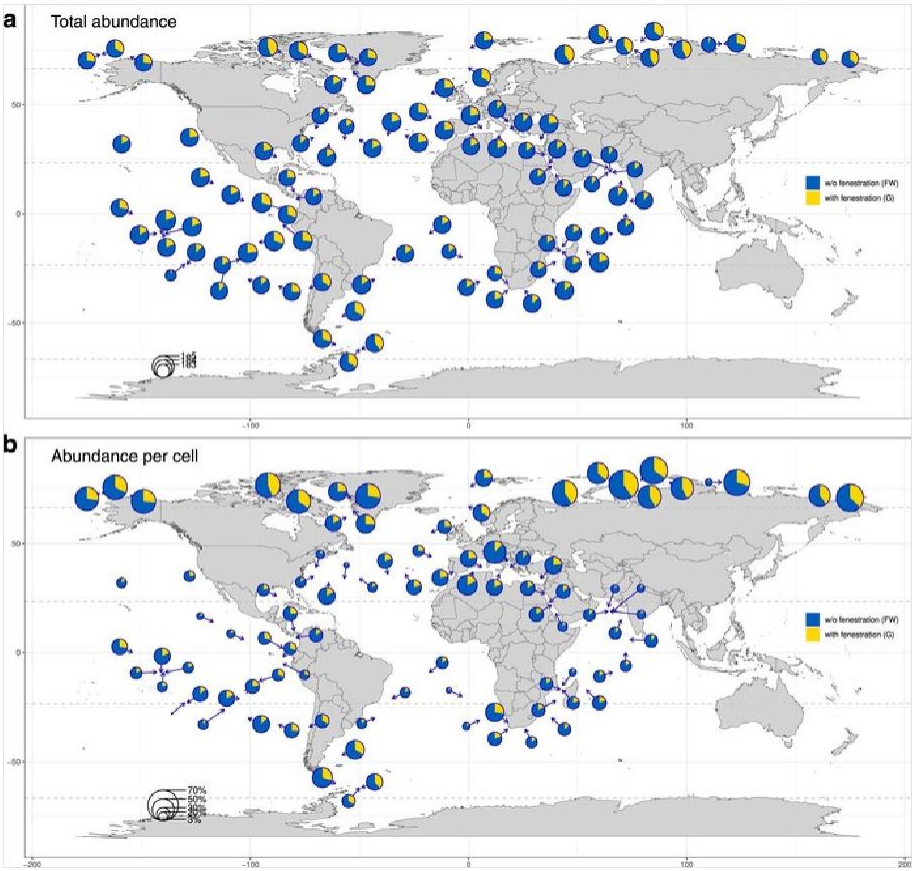


Extended Data Fig. 10

Global distribution of fenestrated XRs and PRs based on the *Tara* Oceans data.

Rhodopsin sequences from the Ocean Microbial Reference Catalog v.2 were classified into PRs and XRs with the canonical DTE motif, and binned into those having bulky residues (Phe or Trp) or Gly at XR position 156. Individual pie charts represent ratios between fenestrated (yellow) and non-fenestrated (blue) PRs and XRs by location and the size of the charts is proportional to the total abundance of DTE PR- and XR-encoding ORFs (**a**) and this abundance relative to the abundance of ten single-copy markers (**b**).

Source data



Supplementary information

Supplementary Information

This file contains Supplementary Figs. 1–2 and Supplementary Tables 1–2.

Reporting Summary

Supplementary Data

Source Data for Supplementary Fig. 2.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 5

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Source Data Extended Data Fig. 10

References

1. Balashov, S. P. et al. Xanthorhodopsin: a proton pump with a light-harvesting carotenoid antenna. *Science* **309**, 2061–2064 (2005).

2. Imasheva, E. S., Balashov, S. P., Choi, A. R., Jung, K.-H. & Lanyi, J. K. Reconstitution of *Gloeobacter violaceus* rhodopsin with a light-harvesting carotenoid antenna. *Biochemistry* **48**, 10948–10955 (2009).

3. Fuhrman, J. A., Schwalbach, M. S. & Stingl, U. Proteorhodopsins: an array of physiological roles? *Nat. Rev. Microbiol.* **6**, 488–494 (2008).

4. Vollmers, J. et al. Poles apart: Arctic and Antarctic *Octadecabacter* strains share high genome plasticity and a new type of xanthorhodopsin. *PLoS ONE* **8**, e63422 (2013).

5. Bertsova, Y. V., Arutyunyan, A. M. & Bogachev, A. V. Na^+-translocating rhodopsin from *Dokdonia* sp. PRO95 does not contain carotenoid antenna. *Biochem. Mosc.* **81**, 414–419 (2016).

6. Misra, R., Eliash, T., Sudo, Y. & Sheves, M. Retinal–salinixanthin interactions in a thermophilic rhodopsin. *J. Phys. Chem. B* **123**, 10–20 (2019).

7. Béjà, O. et al. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**, 1902–1906 (2000).

8. Béjà, O., Spudich, E. N., Spudich, J. L., Leclerc, M. & DeLong, E. F. Proteorhodopsin phototrophy in the ocean. *Nature* **411**, 786–789 (2001).

9. Atamna-Ismaeel, N. et al. Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems. *ISME J.* **2**, 656–662 (2008).

10. Frigaard, N.-U., Martinez, A., Mincer, T. J. & DeLong, E. F. Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* **439**, 847–850 (2006).

11. Finkel, O. M., Béjà, O. & Belkin, S. Global abundance of microbial rhodopsins. *ISME J.* **7**, 448–451 (2013).

12. Gómez-Consarnau, L. et al. Microbial rhodopsins are major contributors to the solar energy captured in the sea. *Sci. Adv.* **5**, eaaw8855 (2019).

13. DeLong, E. F. & Béjà, O. The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *PLoS Biol.* **8**, e1000359 (2010).

14. Munson-McGee, J. H. et al. Decoupling of respiration rates and abundance in marine prokaryoplankton. *Nature* **612**, 764–770 (2022).

15. Wang, W.-W., Sineshchekov, O. A., Spudich, E. N. & Spudich, J. L. Spectroscopic and photochemical characterization of a de ep ocean proteorhodopsin. *J. Biol. Chem* **278**, 33985–33991 (2003).

16. Man, D. Diversification and spectral tuning in marine proteorhodopsins. *EMBO J.* **22**, 1725–1731 (2003).

17. Lanyi, J. K. & Balashov, S. P. in *Halophiles and Hypersaline Environments* (eds. Ventosa, A., Oren, A. & Ma, Y.) 319–340 (Sp ringer, 2011).

18. Balashov, S. P. et al. Reconstitution of *Gloeobacter* rhodopsin with echinenone: role of the 4-keto group. *Biochemistry* **49**, 979 2–9799 (2010).

19. Kopejtka, K. et al. A bacterium from a mountain lake harvests light using both proton-pumping xanthorhodopsins and bacterio chlorophyll-based photosystems. *Proc. Natl Acad. Sci. USA* **119**, e2211018119 (2022).

20. Pushkarev, A. & Béjà, O. Functional metagenomic screen reveals new and diverse microbial rhodopsins. *ISME J* **10**, 2331–23 35 (2016).

21. Pushkarev, A. et al. A distinct abundant group of microbial rhodopsins discovered using functional metagenomics. *Nature* **558** 595–599 (2018).

22. Chazan, A. et al. Diverse heliorhodopsins detected via functional metagenomics in freshwater Actinobacteria, Chloroflexi and Archaea. *Environ. Microbiol.* **24**, 110–121 (2022).

23. Inoue, K. et al. A light-driven sodium ion pump in marine bacteria. *Nat. Commun.* **4**, 1678 (2013).

24. Bhosale, P. & Bernstein, P. S. Microbial xanthophylls. *Appl. Microbiol. Biotechnol.* **68**, 445–455 (2005).

25. Demmig-Adams, B., Polutchko, S. K. & Adams, W. W. Structure–function–environment relationship of the isomers zeaxanthi n and lutein. *Photochem* **2**, 308–325 (2022).

26. *Microbial Carotenoids: Methods and Protocols* (Humana Press, 2018). **AQ36**

27. Ram, S., Mitra, M., Shah, F., Tirkey, S. R. & Mishra, S. Bacteria as an alternate biofactory for carotenoid production: a review of its applications, opportunities and challenges. *J. Funct. Foods* **67**, 103867 (2020).

28. Shibata, M. et al. Oligomeric states of microbial rhodopsins determined by high-speed atomic force microscopy and circular di chroic spectroscopy. *Sci. Rep.* **8**, 8262 (2018).

29. Luecke, H. et al. Crystallographic structure of xanthorhodopsin, the light-driven proton pump with a dual chromophore. *Proc Natl Acad. Sci. USA* **105**, 16561–16565 (2008).

30. Chuon, K. et al. Assembly of natively synthesized dual chromophores into functional actinorhodopsin. *Front. Microbiol.* **12**, 6 52328 (2021).

31. Yoshizawa, S., Kawanabe, A., Ito, H., Kandori, H. & Kogure, K. Diversity and functional analysis of proteorhodopsin in marin e *Flavobacteria Environ. Microbiol.* **14**, 1240–1248 (2012).

32. Ahmed, F. et al. Profiling of carotenoids and antioxidant capacity of microalgae from subtropical coastal and brackish waters. *Food Chem.* **165**, 300–306 (2014).

33. Shihoya, W. et al. Crystal structure of heliorhodopsin. *Nature* **574**, 132–136 (2019).

34. Kishi, K. E. et al. Structural basis for channel conduction in the pump-like channelrhodopsin ChRmine. *Cell* **185**, 672–689.e23 (2022).

35. Balashov, S. P., Imasheva, E. S., Wang, J. M. & Lanyi, J. K. Excitation energy-transfer and the relative orientation of retinal an d carotenoid in xanthorhodopsin. *Biophys. J.* **95**, 2402–2414 (2008).

36. in *Principles of Fluorescence Spectroscopy* (ed. Lakowicz, J. R.) 27–61 (Springer, 2006). **AQ37**

37. Dana, J. et al. Testing the fate of nascent holes in CdSe nanocrystals with sub-10 fs pump–probe spectroscopy. *Nanoscale* **13** 1982–1987 (2021).

38. Polívka, T. et al. Femtosecond carotenoid to retinal energy transfer in xanthorhodopsin. *Biophys. J* **96**, 2268–2277 (2009).

39. Iyer, E. S. S., Gdor, I., Eliash, T., Sheves, M. & Ruhman, S. Efficient femtosecond energy transfer from carotenoid to retinal in *Gloeobacter* rhodopsin–salinixanthin complex. *J. Phys. Chem. B* **119**, 2345–2349 (2015).

40. Doi, S., Tsukamoto, T., Yoshizawa, S. & Sudo, Y. An inhibitory role of Arg-84 in anion channelrhodopsin-2 expressed in *Esch erichia coli Sci. Rep.* **7**, 41879 (2017).

41. Nagiri, C. et al. Crystal structure of human endothelin ETB receptor in complex with peptide inverse agonist IRL2500. *Comm un. Biol.* **2**, 236 (2019).

42. Yamashita, K., Hirata, K. & Yamamoto, M. KAMO: towards automated data processing for microcrystals. *Acta Crystallogr. D Struct. Biol* **74**, 441–449 (2018).

43. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

44. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

45. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

46. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr* **66**, 486–501 (2010).

47. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Cry stallogr.* **68**, 352–367 (2012).

48. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structur e determination. *Nat. Methods* **14**, 290–296 (2017).

49. Rubinstein, J. L. & Brubaker, M. A. Alignment of cryo-EM movies of individual particles by optimization of image translation s. *J. Struct. Biol.* **192**, 188–195 (2015).

50. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM recon struction. *Nat. Methods* **17**, 1214–1221 (2020).

51. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-parti cle electron cryomicroscopy. *J. Mol. Biol* **333**, 721–745 (2003).

52. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr* **60**, 2126–2 132 (2004).

53. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

54. Yamashita, K., Palmer, C. M., Burnley, T. & Murshudov, G. N. Cryo-EM single-particle structure refinement and map calculati on using Servalcat. *Acta Crystallogr. D Struct. Biol.* **77**, 1282–1291 (2021).

55. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).

56. Inoue, K. et al. Exploration of natural red-shifted rhodopsins using a machine learning-based Bayesian experimental design. *C ommun. Biol.* **4**, 362 (2021).

57. Salazar, G. et al. Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *C ell* **179**, 1068–1083.e21 (2019).

58. Chen, I.-M. A. et al. The IMG/M data management and analysis system v.6.0: new tools and advanced capabilities. *Nucleic Acids Res* **49**, D751–D763 (2021).

59. Nayfach, S. et al. A genomic catalog of Earth's microbiomes. *Nat. Biotechnol.* **39**, 499–509 (2021).

60. Sunagawa, S. et al. Metagenomic species profiling using universal phylogenetic marker genes. *Nat. Methods* **10**, 1196–1199 (2013).

61. Wickham, H. in *ggplot2* 189–201 (Springer, 2016). AQ38

62. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).

63. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).

64. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).

65. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).