True-atomic-resolution insights into the structure and functional role of linear chains and low-barrier hydrogen bonds in proteins

Valentin Borshchevskiy, Kirill Kovalev, Ekaterina Round, Rouslan Efremov, Roman Astashkin, Gleb Bourenkov, Dmitry Bratanov, Taras Balandin, Igor Chizhov, Christian Baeken, et al.

True-atomic-resolution insights into the structure and functional role of linear chains and low-barrier hydrogen bonds in proteins

Valentin Borshchevskiy[1,2,3,14], Kirill Kovalev[4,14], Ekaterina Round[5,14], Rouslan Efremov[6,7], Roman Astashkin[3,8], Gleb Bourenkov[4], Dmitry Bratanov[1,2], Taras Balandin[1,2], Igor Chizhov[9], Christian Baeken[1,2], Ivan Gushchin[3], Alexander Kuzmin[3], Alexey Alekseev[3], Andrey Rogachev[3,10], Dieter Willbold[1,2,11], Martin Engelhard[12], Ernst Bamberg[13], Georg Büldt[3] and Valentin Gordeliy[1,2,3,8] ✉

Hydrogen bonds are fundamental to the structure and function of biological macromolecules and have been explored in detail. The chains of hydrogen bonds (CHBs) and low-barrier hydrogen bonds (LBHBs) were proposed to play essential roles in enzyme catalysis and proton transport. However, high-resolution structural data from CHBs and LBHBs is limited. The challenge is that their 'visualization' requires ultrahigh-resolution structures of the ground and functionally important intermediate states to identify proton translocation events and perform their structural assignment. Our true-atomic-resolution structures of the light-driven proton pump bacteriorhodopsin, a model in studies of proton transport, show that CHBs and LBHBs not only serve as proton pathways, but also are indispensable for long-range communications, signaling and proton storage in proteins. The complete picture of CHBs and LBHBs discloses their multifunctional roles in providing protein functions and presents a consistent picture of proton transport and storage resolving long-standing debates and controversies.

The hydrogen bond (H-bond) was discovered more than 110 years ago, with more advanced concepts proposed from the 1920s. The substantial role of H-bonds in biology was uncovered in the second part of the twentieth century and was driven very much by the explanation of the first structures of biological macromolecules: DNA and proteins[1–34].

However, some properties of biological systems and nanomaterials could not be explained on the basis of the existing knowledge of H-bond properties. First of all, the mechanism of anomalous proton transport required new concepts beyond the Grotthuss mechanism. In 1978, linear water chains of H-bonds (CHBs, also named proton wires) were proposed as possible pathways for proton transport[5]. Since the year 2000, the concept of proton wires has been instrumental in the explanation of the mechanisms of proton transport, enzymatic catalysis and signaling in macromolecules such as enzymes, proton transporters, G protein-coupled receptors and several nanomaterials. However, it is mainly based on computer simulations or indirect evidence rather than supported by direct experiments[6–10].

Another concept related to the fundamental mechanisms of H-bonds in biological macromolecules is low-barrier H-bonds (LBHBs), which play an important role in enzymatic catalysis[11–15]. For only a few water-soluble proteins, LBHBs have also been shown structurally[6,16,17]. The lack of direct experimental information on the structure and functional role of CHBs and LBHBs does not enable the understanding of their properties in biomacromolecules. Unfortunately, studies of CHBs and LBHBs are challenging, since their 'visualization' requires simultaneous availability of ultrahigh-resolution structures not only of the ground, but also of the functionally important intermediate states of the protein in the course of its work, identification of all proton translocation events and their assignment to the structural states. This is still a challenge and explains why, even the fundamental principles of bacteriorhodopsin (BR) function, the model in the studies of the mechanisms of proton transport and storage, there is still a long-standing debate. It is important to note that both concepts of CHBs and LBHBs were used to explain the mechanisms of proton transport and excess proton storage[7,14,18].

BR is an ideal model not only to gain information on proton transport mechanisms, but also to obtain invaluable information on the fundamentals of H-bonds, in particular on structural and functional properties of CHBs and LBHBs. Indeed, a light-driven proton pump BR from the archaeon *Halobacterium salinarum* serves as a classical model for studies of these phenomena. BR was the first microbial rhodopsin discovered in 1971 (ref. [19]) and is now one of the most studied membrane proteins. It consists of seven transmembrane α-helices (A–G) connected through short loops and harbors all-*trans* retinal as a chromophore[19,20]. Retinal is covalently

[1]Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany. [2]JuStruct: Jülich Center for Structural Biology, Forschungszentrum Jülich, Jülich, Germany. [3]Research Center for Molecular Mechanisms of Aging and Age-Related Diseases, Moscow Institute of Physics and Technology, Dolgoprudny, Russia. [4]European Molecular Biology Laboratory, Hamburg unit c/o DESY, Hamburg, Germany. [5]European X-ray Free Electron Laser GmbH, Schenefeld, Germany. [6]Structural Biology Brussels, Department of Bioengineering Sciences, Vrije Universiteit Brussel, Brussels, Belgium. [7]Center for Structural Biology, Vlaams Instituut voor Biotechnologie, Brussels, Belgium. [8]Institut de Biologie Structurale J.-P. Ebel (IBS), Université Grenoble Alpes, CEA, CNRS, Grenoble, France. [9]Institute for Biophysical Chemistry, Hannover Medical School, Hannover, Germany. [10]Joint Institute for Nuclear Research, Dubna, Russia. [11]Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany. [12]Department of Structural Biochemistry, Max Planck Institute for Physiology, Dortmund, Germany. [13]Max Planck Institute of Biophysics, Frankfurt am Main, Germany. [14]These authors contributed equally: V. Borshchevskiy, K. Kovalev, E. Round. ✉e-mail: valentin.gordeliy@ibs.fr

bound to K216 in the middle of the helix G via a protonated retinal Schiff base (RSB).

On absorption of a photon, BR undergoes a photocycle with the primary reaction of the retinal chromophore isomerization from an all-*trans* to a 13-*cis* form. Consequently, the protein passes within hundreds of femtoseconds through the I and J, and then within picoseconds through the K; after that, it passes within microseconds to the L, tens of microseconds to the M intermediates and finally returns back to the ground state passing the N and O states in milliseconds[21]. Thus, the retinal isomerization triggers structural and electrostatic alterations in the protein driving several subsequent reactions: deprotonation of the RSB and proton transfer to the primary acceptor D85 and simultaneous proton release from the proton release (storage) group (PRG) to the extracellular bulk on L-to-M transition. On the M-to-N state transition, the RSB is reprotonated from the cytoplasmic side via D96. This change of accessibility switch occurs most likely in a spectroscopically silent transition between two M states, that is M_1 and M_2 (ref. [22]).

The early history of BR reveals how it became one of the most studied membrane proteins. It started from the discovery of purple membranes by Stoeckenius and Rowen in 1967 by electron microscopy[23]. The purple membrane is the first natural membrane to be found as a highly ordered 2D crystal, which allowed the solution of the BR structure by electron diffraction, by Unwin and Henderson[24], to 3.5 Å resolution in-plane. The early work manifested the seven transmembrane α-helix scaffold, a hallmark of this membrane protein family, and the location of water molecules[25], necessary elements of the proton pump[26] by neutron diffraction.

A long-held goal of the research on BR is to find an explanation of the molecular mechanisms of proton transport and storage. To reach this goal, enormous efforts have been made to obtain the structure of the protein by high-resolution X-ray crystallography; however, for almost 30 years, it was not possible to grow crystals of proper crystallographic quality. In parallel, electron microscopy, due to outstanding developments, led to the deciphering of the structure at 3 Å (ref. [20]). This revealed the important features of BR for the first time, but still the details of the protein organization beginning with water molecules, which mediate proton translocation through the Grotthuss mechanism, remained elusive. It was not until 1997, due to the ground-breaking development of a new crystallization method (lipidic cubic phase)[27], that small crystals could diffract to high resolution. Finally, the X-ray structure of BR was obtained[27,28].

Despite the fact that since 1999 about 10,000 genes of microbial rhodopsins have been discovered, only some of them have been characterized, including those featuring new functions (light-driven sodium pumps, channelrhodopsins, inward proton pumps and so on[29–35]). Until now, BR has served as the classic reference in the studies of all of them. Even in one of the first optogenetic studies, BR was transferred into the inner mitochondrial membrane for proton pumping[36]. Studies of BR using different methods (X-ray, electron microscopy, Raman spectroscopy, Fourier-transform infrared spectroscopy (FTIR) and so on) are still ongoing. Since 1997, numerous X-ray structures of the protein in different functional states have been solved by different research groups[37–44]. However, important contradictions between the results from different teams on the structures of the intermediate states of BR and corresponding conclusions on the mechanisms of proton transport and storage have not allowed the creation of a reliable molecular picture of proton storage and translocation mechanisms[45,46]. Moreover, there have been inconsistences between structural and spectroscopic data[47].

It has been suggested that the lack of consensus may have originated from differences in crystallization conditions, insufficient resolution of the data, low occupancy of the intermediate states, the presence of merohedral twinning of the best diffracted and most commonly used $P6_3$ crystals and also X-ray radiation-induced changes in protein structure[21,48,49]. To resolve the contradictions, crystallographic data with an atomic resolution, showing low or absent twinning, collected with an ultra-low dose of X-ray radiation and with sufficiently high occupancies of the intermediate states, are required. There have already been several attempts to accomplish these goals, however, none of these efforts satisfied all four criteria simultaneously[21].

One new attempt applied time-resolved crystallography to X-ray free-electron lasers[50–52]. However, the best resolution was just 1.5 Å for the ground state of BR, while being only 1.9 Å for the K and M states and 2.1 Å for the L state. In addition, the emerging time-resolved serial femtosecond crystallography experiments still may suffer from experimental and data interpretation controversies. For instance, it has been pointed out that the pump laser power is far too high compared to what could be expected from the spectroscopic data. It may result in a multiphoton excitation regime[52] followed by distorted photocycle and protein photo-degradation. It has been suggested that it is manifested by negative electron densities appearing randomly at methionine residues in difference maps of both time-resolved serial femtosecond crystallography studies of BR and protein quakes. On top of a low fraction of excited proteins (in some cases about 10%), the analysis is further complicated by the principal inability to obtain structure factors with a single pure intermediate state.

From studies of other proteins, such as myoglobin[53], it became evident that often those structural elements that are important for function require ultrahigh-resolution data. This approach was applied to the proton pathway of BR[54]. However, a resolution of 1.3 Å was only achieved for the ground state and with considerably twinned crystals. Here we present atomic structures of the ground, L and M states, key intermediates of the BR photocycle, solved at 1.05–1.2 Å resolution and the structure of the K state solved at 1.53 Å. The data were collected at a synchrotron source and are both free of twinning and X-ray radiation-induced damage problems with occupancies of the intermediate states in the range of 15 to 50%. Our results demonstrate fundamental principles of proton storage and a CHB-based molecular mechanism of proton transport in BR and extend the existing knowledge about the structure and function of H-bonds.

Results and discussion

True-atomic-resolution structural data collection procedures. The crystals of $P6_3$ symmetry of BR were grown using the in meso method similar to our previous studies[55,56]. To eliminate or reduce the twin fraction, we used an approach described in ref. [48]. To avoid X-ray radiation-induced damage of the crystals for the data collection, we used the protocols reported in refs. [49,57].

Crystallographic data on BR were collected at the European Synchrotron Radiation Facility (ESRF, Grenoble). To obtain the best structural data quality and check reproducibility, we used several high-quality crystals for each functional state. For instance, we solved the structure of BR ground state with the highest resolution of 1.05 Å. However, the dataset had a moderate twinning (Table 1). Another dataset without twinning was collected, and the structure was solved at 1.22 Å (Table 1). The comparison of electron densities of both datasets indicated the identical nature of the final models. Following a similar approach, we obtained the highest-quality structures of the K, L and M intermediate states of BR at 1.53, 1.2 and 1.22 Å, respectively. The intermediates were cryogenically trapped in the crystals using the previously described methods[58–60]. The exact procedures of the accumulation and spectroscopic characterization of each intermediate state in crystals are described in the Methods section and Extended Data Fig. 1. The structures of the intermediate states were determined using difference electron density ($F_{\text{obs light}} - F_{\text{obs dark}}$ difference Fourier electron-density maps) (Extended Data Figs. 2 and 3).

Table 1 | Data collection and refinement statistics of BR

	Ground state 1 PDB ID 7Z09	Ground state 2 PDB ID 7Z0A	K state PDB ID 7Z0C	L state PDB ID 7Z0D	M state PDB ID 7Z0E
Data collection					
Space group	$P6_3$	$P6_3$	$P6_3$	$P6_3$	$P6_3$
Cell dimensions					
a, *b*, *c* (Å)	60.73, 60.73, 110.27	60.97, 60.97, 109.56	60.99, 60.99, 110.08	60.92, 60.92, 110.18	60.89, 60.89, 109.23
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	35–1.05 (1.08–1.05)[a]	35–1.22 (1.25–1.22)[a]	50–1.53 (1.57–1.53)[a]	35–1.20 (1.23–1.20)[a]	35–1.22 (1.25–1.22)[a]
R_{merge} (%)	8.7 (291.5)	4.1 (250.9)	3.2 (228.5)	5.7 (192.6)	6.6 (220.9)
$I/\sigma I$	9.36 (0.56)	24.6 (0.3)	21.9 (0.3)	20.8 (0.4)	13.8 (0.4)
Completeness (%)	99.4 (94.8)	91.2 (42.4)	86.9 (39.2)	96.7 (73.7)	98.1 (86.1)
Unique reflections	106,346 (7,436)	63,006 (2,159)	30,368 (1,013)	69,741 (3,901)	66,766 (4,327)
Redundancy	6.1 (3.7)	6.8 (1.4)	3.3 (1.3)	6.2 (1.6)	6.4 (2.3)
$CC_{1/2}$ (%)	99.9 (18.0)	100.0 (15.0)	100.0 (30.0)	100.0 (17.7)	99.9 (17.3)
Twin fraction (%)	40	0	0	5	32
Occupancy of the intermediate (%)			20	25	45
Refinement					
Resolution (Å)	30.13–1.05	29.37–1.22	38.11–1.53	29.36–1.20	30.44–1.22
No. reflections	1,05,747	62,233	29,725	68,240	66,454
R_{work} / R_{free}	17.5 / 18.9	14.1 / 16.7	15.0 / 17.1	15.5 / 17.4	18.2 / 19.5
No. atoms					
Protein	1,931	1,931	1,969	2,447	2,835
Retinal	20	20	40	40	40
Water	88	88	89	98	88
Lipid	257	257	257	252	278
B factors					
Protein	20	26	29	23	22
Retinal	13	19	22	16	16
Water	44	47	49	42	43
Lipid	42	51	54	48	47
R.m.s. deviations					
Bond lengths (Å)	0.008	0.006	0.014	0.011	0.011
Bond angles (°)	1.257	0.863	1.758	1.433	1.792

[a]Values in parentheses are for the highest-resolution shell.

The statistics of the data treatment and structure refinement are presented in Table 1. The quality of the data and corresponding structures is also confirmed by the number of collected unique reflections, resolved water molecules both inside the protein and on its surface, and numerous lipid fragments, almost entirely covering the hydrophobic region of BR (Extended Data Fig. 4). These extensive data considerably exceed what has been published previously[45,50,51]. For example, the number of collected unique reflections for the L state in the present case (69,741) surpasses those of the best low-twinning data[50] by a factor of five. The residues 5–234 are resolved, including the EF-loop that is absent in the most commonly used model of BR[37] (Protein Data Bank (PDB) ID 1C3W). Taken together, the quality of the data allowed reliable identification of functionally essential water molecules, conformations of the key molecular groups of the protein and the length of H-bonds. In Supplementary Notes 1–4 we describe the key newly identified features of BR structure in the ground state (Fig. 1) and demonstrate structural rearrangements in the central, cytoplasmic and extracellular regions of the protein during a photocycle (Figs. 2–4 and Extended Data Figs. 5–8). Below, we discuss BR mechanisms in parallel with functional properties of CHBs and LBHBs as they follow from our experimental data and in the context of the current knowledge. It should be noted that LBHBs are assigned to very short (2.3–2.5 Å) H-bonds found in the structures[61]. More details on the LBHB tracing are given in the Methods section of the paper.

Mechanism of proton storage by PRG. The structural data presented here indicate a molecular mechanism of proton storage at the PRG of BR (Figs. 4 and 5a). The excess proton is stored in a delocalized form. It is manifested by the presence of several short LBHBs in the PRG and also by the multiple conformations of the PRG in the ground and K states (Figs. 4c and 5a). The PRG is located close to the hydrophobic/hydrophilic membrane boundary and is more dynamic than the central part of the protein, which might result in the existence of several local minima in its energy landscape, corresponding to different quasi-stable conformations of the PRG.

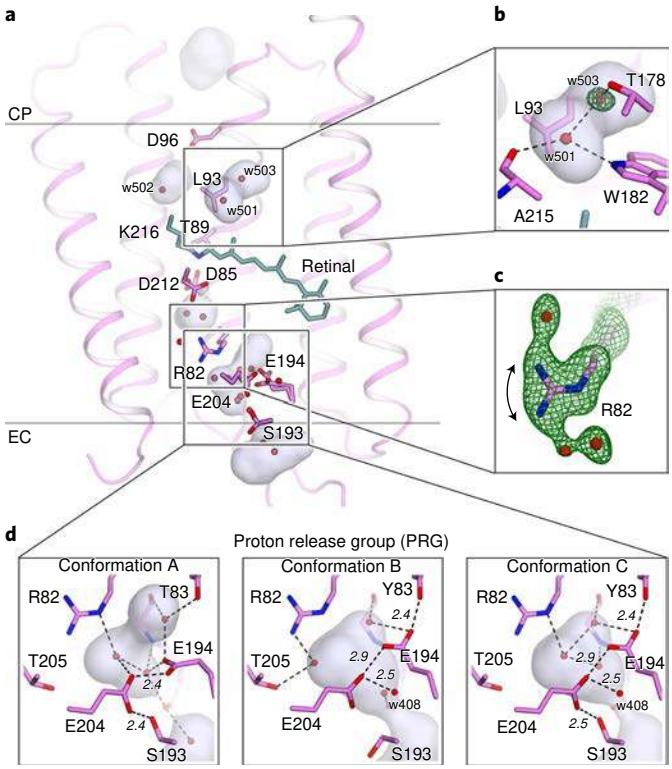


Fig. 1 | True-atomic-resolution structure of BR in the ground state. a, Side view of BR in the ground state. Retinal cofactor is colored teal. Cavities were calculated using HOLLOW[80] and are shown with the grey-blue surfaces. Hydrophobic/hydrophilic membrane core boundaries were calculated using the Positioning of Proteins in Membrane server[81] and are shown with gray horizontal lines. **b**, Region of the additional water molecule w503. A Polder[82] map omitting the w503 molecule is shown with green mesh and is contoured at the level of 3σ. **c**, A Polder[82] map omitting R82 and nearby water molecules is shown with red mesh and is contoured at the level of 5σ demonstrate R82 dynamics in the ground state of BR. **d**, Three conformations of the PRG. The lengths of the key very short H-bonds are shown with italic numbers and are in Å (italic numbers). Additional water molecule w408 is labeled.

The possibility of multiple conformations of the PRG in the ground state of BR was first proposed in ref. [37]. However, structural resolution of the data was not sufficient to resolve different conformations. The double conformation of the PRG was suggested in the 1.3 Å resolution structure of BR in the ground state[54]. This structure was obtained with highly twinned crystals and the resolution of the data was not sufficient to resolve an additional conformation, which, in turn, could corroborate the two conformations presented in ref. [54]. The data presented in our work show the triple conformation of the PRG in the ground and K states of BR (Fig. 4a). We should note that the known conformations of the PRG in the available structures of BR at lower resolution should be considered as an average of its nonresolved multiple states evidenced by the diffuse electron-density maps in the region[37].

The major conformation (approximately 50% occupancy) of the PRG is characterized by a very short H-bond between E194 and E204 (2.4 Å) indicating that the proton is delocalized (shared) between the residues (Figs. 4a and 5a). Such a short H-bond between E194 and E204 was also observed recently in detailed (quantum mechanic/molecular mechanic) simulations[18]. In the other two conformations (approximately 25% and 25% occupancies) E194 and E204 also interact directly with each other; however, the distance between them is larger (2.9 Å), which likely means that the proton is localized on one of the residues. It is also possible that in each of the two minor conformations of the PRG the proton is transiently stored at different glutamates or shared (delocalized) between the glutamates and neighboring amino acids connected with E194 and E204 by short H-bonds, such as Y83 or S193 (Fig. 5a). Exactly the same triple conformation is found in the K state of BR. However, only one conformation of PRG is found in the L state. This single conformation also corresponds to that of the delocalized proton in the very short H-bond (2.3 Å) between E194 and E204 (Fig. 5a). Two other short H-bonds (between R82 and E194 (2.4 Å) and between E204 and w408′ (2.5 Å)) are found in the structure of the PRG. This may mean that in the L state the excess proton, just before it is released, is delocalized within a short proton wire R82-E194–E204-w408′.

Thus, in the ground, K and L states PRG in all conformations is characterized by several very short H-bonds, which is characteristic of delocalized protons and e^- holes[11–13] (Figs. 4 and 5a). This structure-based suggestion is supported by the experimentally observed continuous infrared band in the range of 1,800–2,000 cm^{-1} (ref. [62]). Our structure-based suggestion is also strongly supported by the fact that the short H-bonds in PRG are not observed in the M state, which correlates with the absence of the continuum absorbance band described by K. Gerwert and coworkers[62]. Moreover, time-resolved step-scan FTIR data showed that in the L-to-M state transition the proton is released from the PRG from its delocalized form, but not from a single residue such as E194 or E204 (refs. [63,64]). It also supports the hypothesis that the stored proton is delocalized over the entire PRG including some water molecules[14].

Based on FTIR data[37], it was previously suggested that the proton is stored at the PRG in a cluster of water molecules rather than at the carboxylic acid residues[62,65]. Our structural data show that the proton is always (at least in the major fraction of PRG in the ground, K and in a single conformation in the L state) at the E194–E204 pair, stabilized by the surrounding residues such as R82, Y83, S193, T205 and water molecules.

We suggest that the biological functional role of the proton delocalization in the PRG is the maintenance of its high proton affinity. Indeed, it has been argued that to store excess protons close to the bulk with high proton concentrations, it is necessary to keep the high pK_a of the PRG that cannot be provided by negatively charged amino acids separately or Eigen or Zundel cations[66]. Vice versa, a high proton affinity can be provided by LBHBs between the amino acids where the proton is stored in a delocalized form[66].

Mechanism of proton release from the PRG. On the L-to-M state transition, the direct connection between E194 and E204 is lost and the proton is released from the PRG (Fig. 5a). However, an important question concerns the trigger of these events. It is commonly accepted that such a trigger is the deprotonation of the RSB to D85 and the corresponding flip of R82 (refs. [46,50]). However, the role of R82 in the proton translocation by BR is a long-standing problem. About a half of the known structures of the L state demonstrate the presence of a flip of R82 toward the PRG, while the other do not[45,50]. Our true-atomic-resolution structure of the L state shows that the flip of R82 itself does not trigger the proton release from the PRG. Indeed, the PRG remains protonated despite the partially flipped R82. Moreover, R82 is already mobile in the ground state of BR (Fig. 1). A similar feature has been recently demonstrated for the corresponding arginine of archaerhodopsin-3 (ref. [67]).

Based on true-atomic-resolution structural data presented here, we suggest the following mechanism of proton release from the PRG. During the K-to-L state transition, dehydration of the extracellular part of BR results in the R82 flip toward the PRG. At this step, a single CHB is formed connecting D85 to the PRG (Fig. 5b). D85 is negatively charged in the L state; the presence of the negative charge at one side of the CHB results in the shift of the positive charge at

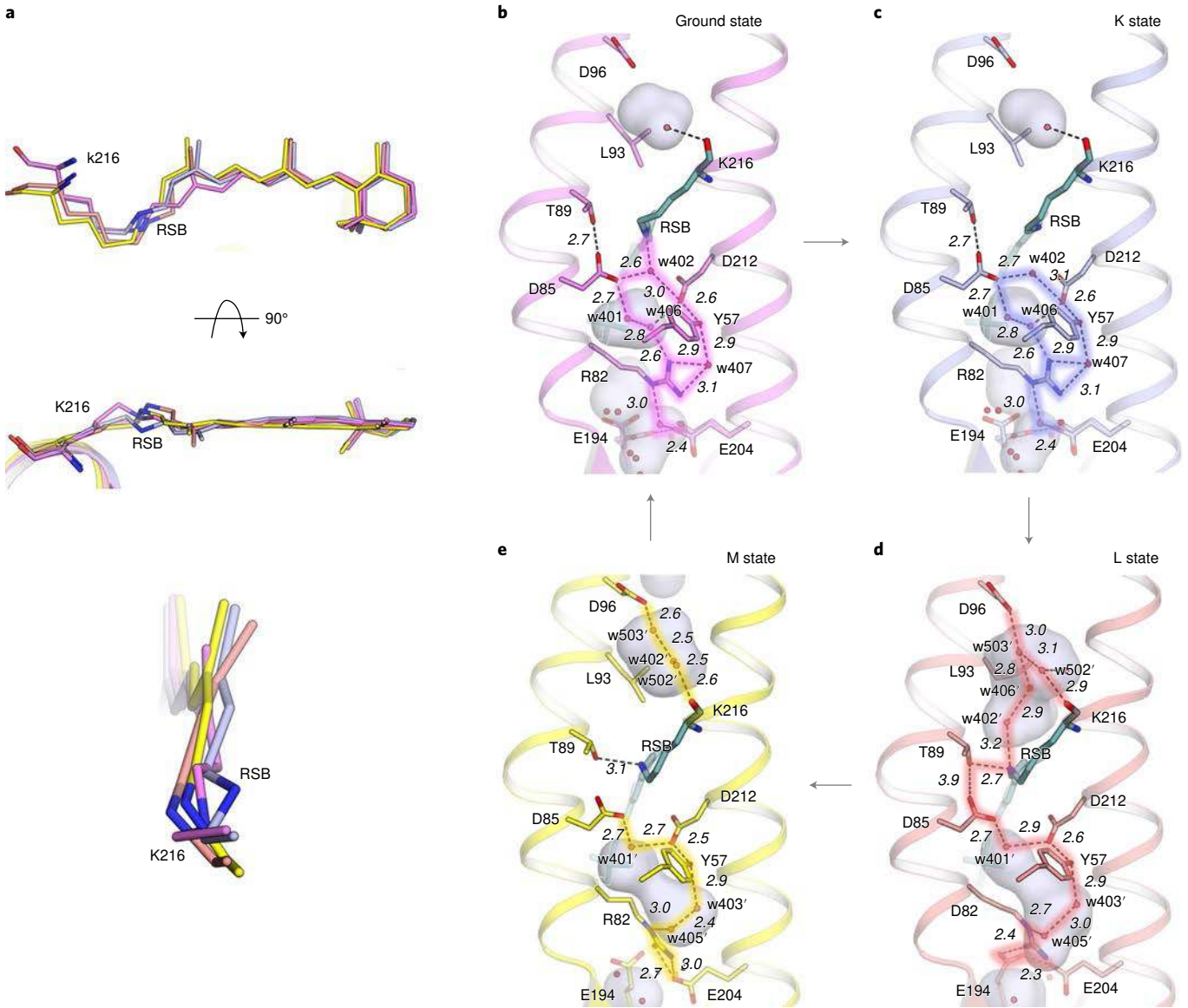


Fig. 2 | BR evolution during the photocycle. a, Evolution of the retinal cofactor and the RSB during the photocycle. The ground, K, L and M states are shown with violet, blue, salmon and yellow colors, respectively. **b**, Central and cytoplasmic region of BR in the ground state (violet). **c**, Central and cytoplasmic region of BR in the K state (blue). **d**, Central and cytoplasmic region of BR in the L state (salmon). **e**, Central and cytoplasmic region of BR in the M state (yellow). Cavities are calculated using HOLLOW[81] and are shown with grey-blue surfaces. H-bonds are shown with black dashed lines, and corresponding distances are given in Å (italic numbers). Key CHBs (proton wires) are highlighted violet (in the ground state), blue (in the K state), salmon (in the L state) and yellow (in the M state). The retinal cofactor is colored teal in **b**–**e**. Water molecules are shown with red spheres. Helices A and B are shown as a cartoon representation.

R82 toward the NE atom. On the L-to-M transition, D85 accepts the proton from the RSB and becomes neutral. Consequently, the polarization along the CHB is modified leading to the redistribution of the positive charge at R82. Thus, the positive charge shifts toward NH1 and NH2 atoms of R82. The latter event likely causes the proton release from the PRG and following formation of the salt bridge between R82 and negatively charged in the M state E204, similar to that known from the visual rhodopsin and some other G protein-coupled receptors as an ionic lock[68–70] (Fig. 5a). The ionic lock stabilizes the PRG in the locked conformation preventing proton backflow from the extracellular bulk (Fig. 5a). It is known that although protonation of D85 and proton release take place at spatially distinct sites, these two reactions appear as almost simultaneous events with a very small temporal delay between them[71]. This fact is directly explained within the above-proposed mechanism.

To summarize, we claim that proton release from the PRG is triggered by the D85 protonation on the L-to-M transition. The long-range interactions between D85 and E194–E204 pair are mediated by the single linear CHB preformed already in the L state after R82 flip toward the PRG (Fig. 5b). This shows that CHBs play an essential role in long-distance signaling between the key functional groups in BR. Thus, in contrast to the most commonly accepted hypothesis that the flip of R82 triggers the proton release from BR, we suggest that the proton release is coupled with the RSB deprotonation to D85 by the molecular mechanism resembling to some extent that of the Newton's cradle (Fig. 5c). In this approximation, the water molecules and key amino acid residues forming the linear CHB at the extracellular side of BR serve as transmitters of the polarization switches from D85 to E194–E204 pair. Notably, similar mechanisms based on switches within proton wires in the active

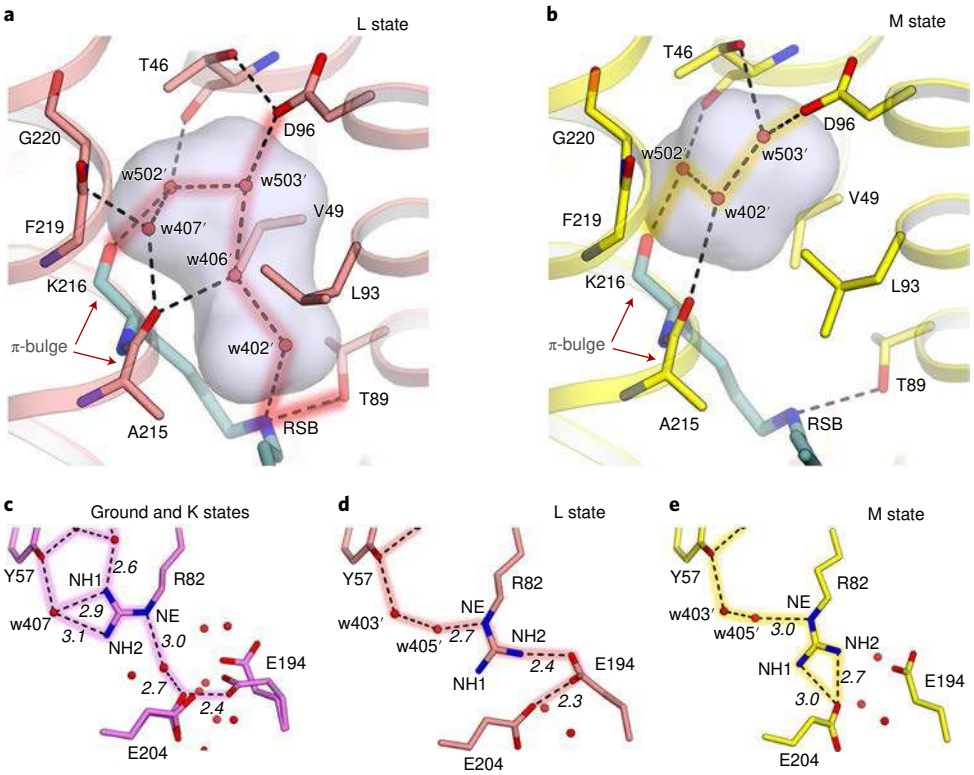


Fig. 3 | Key functionally important switches in BR. a, CHBs (proton wires) in the cytoplasmic inner part of BR in the L state. Three water molecules (w402′, w406′ and w503′) mediate linear CHB between the RSB and D96. Two water molecules (w407′ and w502′) stabilize the enlarged π-helical region (indicated with red arrows) of helix G in the L state. Cavities were calculated using HOLLOW[81] and are shown with grey-blue surfaces. **b**, CHB in the cytoplasmic inner part of BR in the M state. **c**, Region of the R82 residue in the ground and K states. **d**, Region of the R82 residue in the L state. **e**, Region of the R82 residue in the M state. H-bonds are shown with black dashed lines, and corresponding key distances are given in Å (italic numbers). Water molecules are shown with red spheres. H-bond proton wires are highlighted with violet, red and yellow for the ground, L and M state, respectively.

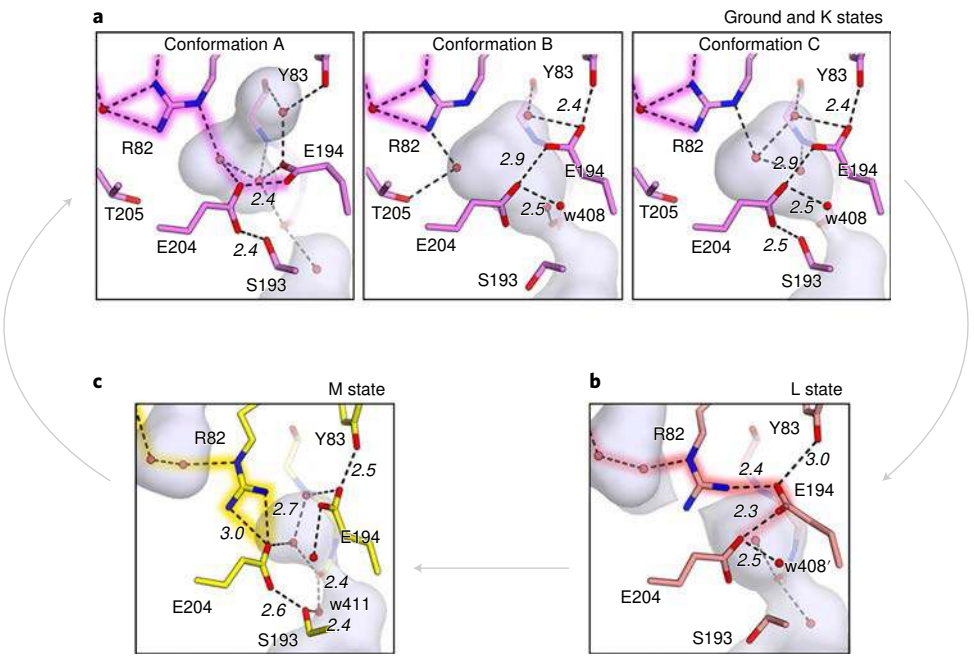


Fig. 4 | PRG evolution during the BR photocycle. a, Triple conformation of the PRG in the ground and K states of the BR photocycle. **b**, Single conformation of the PRG with an extremely short H-bond between E194 and E204 in the L state of the BR photocycle. **c**, Single conformation of the PRG after proton release in the M state of the BR photocycle. Cavities are calculated using HOLLOW[81] and are shown with grey-blue surfaces. H-bonds are shown with black dashed lines, and corresponding key distances are given in Å (italic numbers). Water molecules are shown with red spheres. The termini of the CHBs connecting D85 and the PRG are highlighted violet (in the ground and K states), salmon (in the L state) and yellow (in the M state).

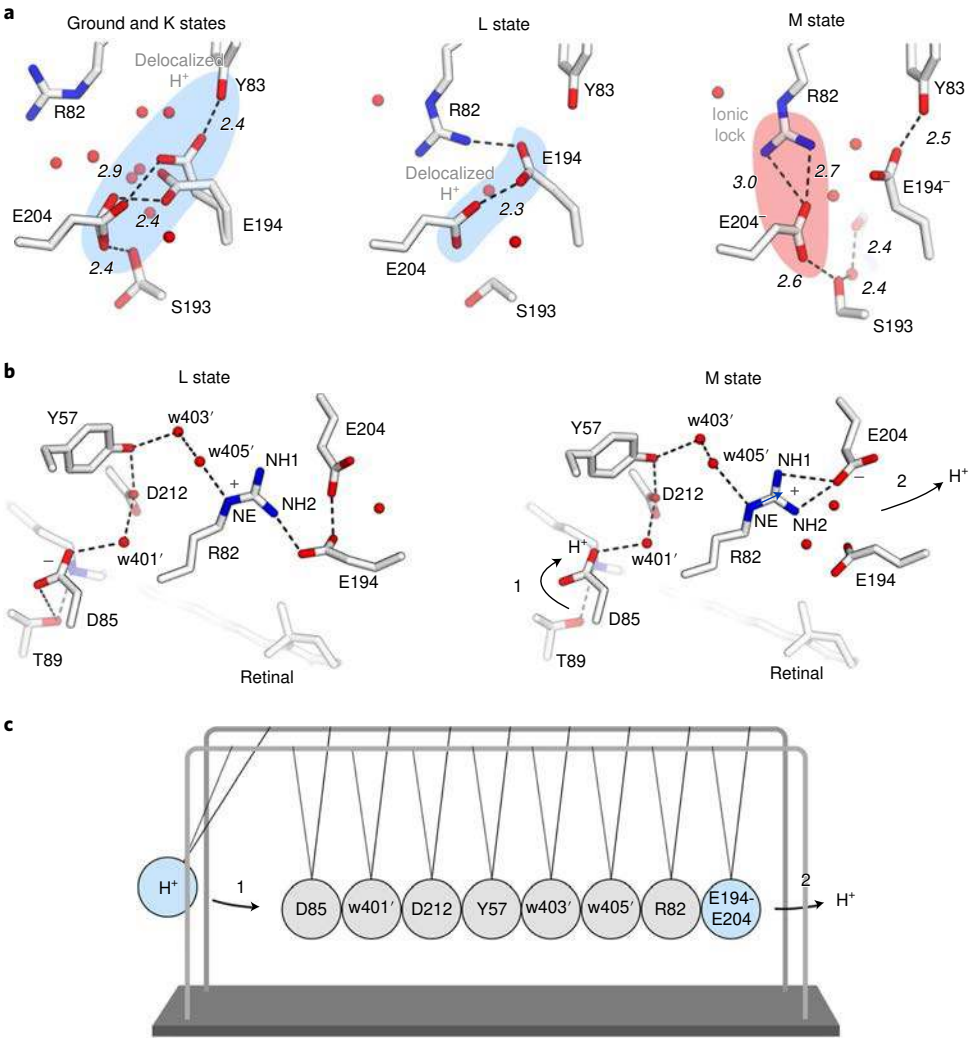


Fig. 5 | Mechanisms of proton storage and release in BR. a, The PRG of BR in the ground, K, L and M states. Proton delocalization in the ground, K and L states is indicated in light blue. The ionic lock in the M state of BR is indicated in red. The key H-bonds are shown with black dashed lines, and corresponding key distances are given in Å (italic numbers). **b**, CHB connecting D85 and PRG in the L and M states. The CHB is formed in the L state. The protonation of D85 (black arrow 1) leads to proton release from PRG (black arrow 2). H-bonds providing the long-range interactions between D85 and PRG are shown with black dashed lines. Water molecules are shown with red spheres. The blue plus sign indicates the positive charge at R82. The blue arrow indicates the redistribution of the positive charge at R82 from NE atom toward NH1 and NH2 atoms on the L-to-M state transition. Red minuses indicate the negative charges at D85 (left panel) and E204 (right panel). **c**, A Newton's cradle schematic model of the proton release from BR. Protonation of D85 (black arrow 1) leads to the proton release from the E194–E204 pair (black arrow 2). The signal is transmitted via the CHB, shown in detail in **b**.

centers have been shown for other proteins, such as human transketolase and pyruvate oxidase from *Lactobacillus plantarum*[6] and glycoside hydrolase from *Phanerochaete chrysosporium*[72].

Although the structure of the M state does not show the signs of the presence of a proton in the PRG, two very short H-bonds (2.4 Å each) are observed near S193 and P77 at the surface of the protein (Fig. 5a). Staying in line with the concept that such short H-bonds are LBHBs with delocalized proton, we speculate that on the L-to-M state transition the PRG proton is not released directly to the bulk but is first trapped for some time at the surface of the protein in delocalized form. Despite this hypothesis being in line with the experimental fact that the released proton stays roughly 1 ms at the purple membrane surface before it appears in the bulk[73], additional studies are necessary to verify our hypotheses on one of possible molecular natures of the retarded surface-to-bulk proton transfer.

CHBs modulate proton affinity of key molecular groups in BR. Our true-atomic-resolution data demonstrate switching on and off of the CHBs in the cytoplasmic/extracellular parts of the protein, respectively, during the photocycle (Fig. 2). These switches are mediated by the dehydration of the extracellular part and synchronous hydration of the cytoplasmic part. Thus, in addition to the above-described signaling role of CHBs, our structures show that CHBs play another two key roles in proton translocation. First, the CHBs between functional molecular groups provide the pathways for proton translocation. Second, the number of CHBs and corresponding H-bonds around a rechargeable functional group directly influences its proton affinity. The affinities of the groups of the both ends of the CHBs determine the direction of proton translocation. Therefore, we suggest that the switches on and off of the CHBs provide the unidirectional proton translocation of BR by regulating proton affinities of the key functional groups of the rhodopsin. The mechanisms of these CHBs remodulations are described below.

First, the two CHBs connect D85 to R82 in the ground and the K states and, together with the direct H-bond of D85 to T89, maintain low proton affinity of the proton acceptor. Second, on the K-to-L

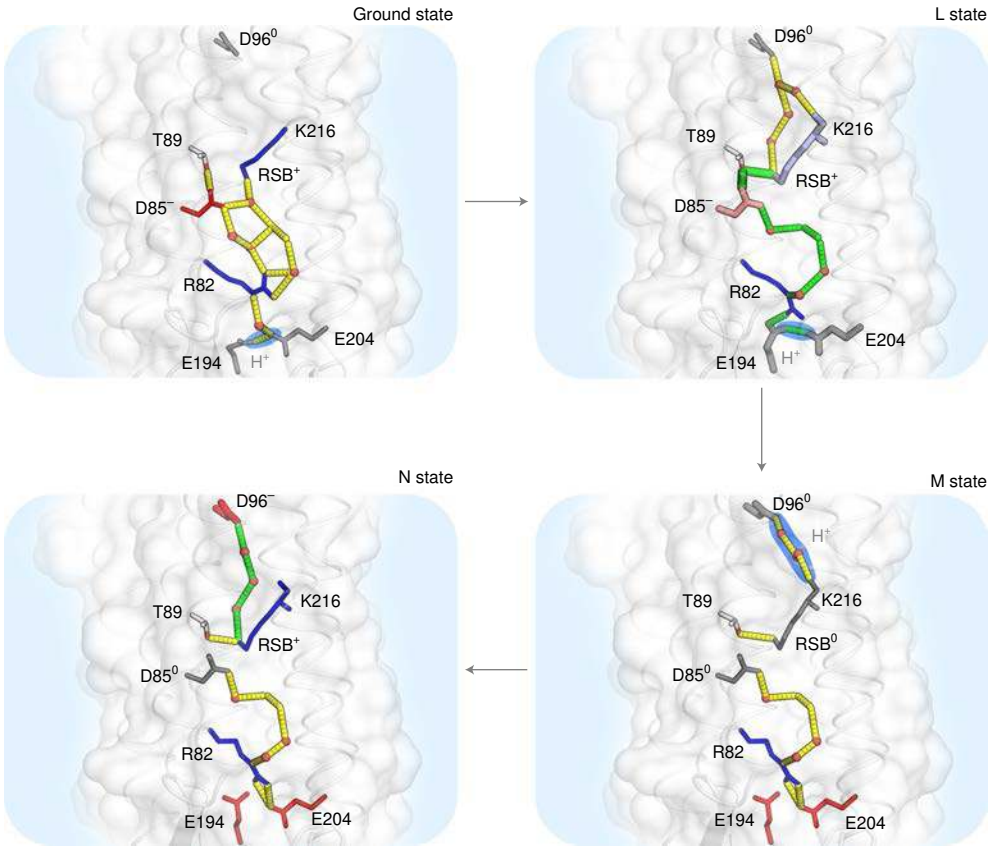


Fig. 6 | Schematic representation of the mechanism of proton storage and vectorial translocation mediated by CHB switches in the course of the BR photocycle. Key rechargeable groups of BR are shown as sticks. CHBs are shown as yellow and green lines. Green color indicates the regions of the continuous CHB actively involved at a certain step of the photocycle. Blue color indicates the positive charge at the key residues and the RSB. Blue highlighting indicates delocalized proton in LBHBs. Red color indicates the negative charge at the key residues. Gray color indicates the neutral charge at the residues. Water molecules mediating the CHBs are shown as red spheres. Light-blue and light-red colors of D85 and RSB, respectively, in the L state indicate the change of pK_a of the residues and the increased (compared to the ground state) probability of their recharge. Blue areas indicate the hydrophobic region of the lipid membrane surrounding the protein.

state transition, reorientation of N-H$^+$ of the RSB to the cytoplasmic side leads to the considerable dehydration of the region between the retinal and R82. Both D85 and R82 reorient toward the PRG already in the L state. Thus, instead of two above-mentioned CHBs, only one connects D85 to R82 at this step. The H-bond between D85 and T89 also becomes weaker. All this results in the increase of the proton affinity of D85 already in the L state.

At the same time, the cytoplasmic part of the protein becomes hydrated in the L state. The bend of the helix G near π-bulge at A215 residue (Supplementary Note 5) and the flip of the L93 side chain create a large cavity at the cytoplasmic side in the region between the RSB and its proton donor D96. Five water molecules are found in the cavity. Three of these water molecules mediate a linear CHB connecting the RSB and D96 in accordance with FTIR data[74]. The CHB stabilizes the protonated RSB oriented to the hydrophobic environment following the retinal isomerization. The RSB is additionally stabilized by the H-bond with T89. This bond and the CHB propagating to D96 help to keep proton affinity of the RSB at a sufficiently high level preventing its deprotonation in the L state. Thus, the switching on of the CHBs mediated by water molecules in this region is the main functional event triggered by the reorientation of N-H$^+$ moiety of the RSB from the extracellular to the cytoplasmic part on the K-to-L state transition due to the isomerization of retinal.

Third, during the L-to-M transition, the linear CHB between the RSB and D96 disappears. We suggest that to switch off of this CHB leads to the decrease in proton affinity of the RSB. Indeed, at this step the RSB is only stabilized by the T89, which is insufficient to maintain its protonation. This, together with the synchronous increase of pK_a of D85, leads to the proton transfer from the RSB to D85 through a short proton wire, RSB–T89–D85, observed in the L state. The mechanism of proton transfer from the RSB to D85 on the L-to-M state transition is a key question. Indeed, while the H-bond between the RSB and T89 is present in both the L and the M states, the connection between D85 and T89 is absent in the M state (Extended Data Fig. 9). In contrast, in the L state, the distance between T89 and D85 is 3.9 Å, which, in accordance with Jeffrey's classification of H-bonds in the range of 3.2–4.0 Å, should be considered as a 'weak, electrostatic' H-bond[4]. We have to note that the 3.9 Å is the average length. Due to thermal fluctuation of molecular groups of the protein, the distance between T89 and D85 fluctuates, which should facilitate the proton transfer. Indeed, the L state decays to the M state within tens of microseconds, which is a long time, compared to the characteristic time of thermal fluctuation of molecular groups in proteins, to be sufficient for an efficient proton transfer[75]. Therefore, we suggest that proton transfer from the RSB to D85 proceeds through the H-bond pathway from the RSB to T89 and from T89 further to the primary proton acceptor D85 through the weak H-bond between these amino acids.

Last, in the M state, the CHB connecting D96 and carbonyl oxygen of K216 is preserved but modified compared to the L state

(Fig. 4). We expect that this CHB is switched toward the deprotonated RSB to provide a pathway for its reprotonation from D96 on the M-to-N state transition.

Molecular mechanism of proton transport in BR. The proposed mechanism of proton translocation against electrochemical gradient by BR based on CHBs and their switches is summarized in Fig. 6. A global CHB passes through highly conservative residues D96–T89–D85–Y57–D212–R82–E194–E204 and involves highly conservative A215 (Extended Data Fig. 10). All these residues are known to be critical for proton pumping. This fact also supports the key role of the proton wires in BR function.

The photon absorption by the retinal triggers its isomerization from the all-*trans* to a 13-*cis* conformation with the rise of the K state and the corresponding slight relocation of the RSB. The H-bond of the RSB to the w402 is already broken in the K state. Nevertheless, the two CHBs connecting D85 and R82 in the ground state are preserved in the K state. In the ground and K states the cytoplasmic part of BR lacks any CHB and is inactive. On the K-to-L transition, two linear water-mediated CHBs appear and connect D96 to the RSB and carbonyl oxygen of K216 (Fig. 6). Thus, the activation of this region and switching on of the CHBs maintain the protonation of the RSB after its relocation to the hydrophobic environment. Simultaneously, the number of CHBs at the extracellular part of BR stabilizing D85 decreases, which makes the presence of the charge at D85 energetically less favorable (Fig. 6). Consequently, the further evolution of the retinal cofactor to a fully relaxed 13-*cis* configuration[76], together with the increasing proton affinity of D85, leads to its protonation from the RSB on the M state formation when the CHB connecting the RSB to D96 disappears. The CHB in the cytoplasmic part of BR connecting D96 to K216 is preserved in the M state (Fig. 6).

Dehydration of the region between the RSB and charged R82 with the L state formation leads to the flip of the latter amino acid toward the PRG. The single long CHB connecting RSB–T89–D85 region with the PRG is formed in the L state substituting for the two CHBs found in the ground and K states (Fig. 6). The flip of R82 is not yet solely sufficient to release the delocalized proton stored in LBHB between E194 and E204 in the L state. The proton is released from the PRG on D85 protonation from the RSB via T89. The long-range communication between D85 and the PRG is provided by the CHB in the extracellular part of BR. We propose that the proton released from the PRG is transiently stored at the protein surface in LBHBs near S193 in the M state. R82 and E204 form an ionic lock preventing the proton backflow from the extracellular space.

Finally, in the N state, the CHB similar to that found in the L state is formed in the cytoplasmic region between the RSB and D96. It serves as a pathway for RSB reprotonation from D96. Unfortunately, high-resolution data on the N and O intermediates of BR revealing all water molecules and CHBs are not yet available. However, a recent structure of the N state solved at 2.6 Å resolution and low occupancy of the intermediate is encouraging[77]. Nevertheless, ultrahigh-resolution structures of the late intermediates of BR photocycle free from above-discussed crystallographic problems would complete a detailed molecular picture of vectorial proton transport from proton release to proton uptake. High-resolution structures of the sub-M states[78] (such as M_1 discussed in ref. [79]) free of the crystallographic problems would also be desirable adding currently missing details regarding proton transfer from the RSB to the primary proton acceptor D85.

Concluding remarks

The results of this work settle the long-standing debate on proton-pumping mechanisms enabling the creation of a comprehensive and consistent picture of the molecular mechanisms of proton storage and vectorial transport. Moreover, they extend the existing knowledge on the structure and function of H-bonds, namely the data experimentally demonstrate the multifunctional structure-functional role of the linear CHBs in proteins.

Our data show that the chains serve as: (1) proton transport pathways; (2) modulators of pK_a values of key functional molecular groups and (3) signal pathways providing long-distance functionally important communications between key amino acid residues. It is shown experimentally that a single chain of H-bonds solely mediated by water molecules provides a pathway for proton transport through a large hydrophobic region of the protein. In addition, we demonstrate that a chain of H-bonds can be mediated solely by water molecules and can completely consist of LBHBs. The data also provide direct experimental support for a hypothesis that LBHBs between glutamic acid residues may serve as excess proton storage and release group.

Online content

Any methods, additional references, Nature Research 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/s41594-022-00762-2.

Received: 23 November 2021; Accepted: 14 March 2022;
Published online: 28 April 2022

References

1. Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: two hydrogen-bonded helical configur tions of the polypeptide chain. *Proc. Natl Acad. Sci. USA* **37**, 205–211 (1951).
2. Pauling, L. & Corey, R. B. The pleated sheet, a new layer configur tion of polypeptide chains. *Proc. Natl Acad. Sci. USA* **37**, 251–256 (1951).
3. Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
4. Jeffrey, G. A. J. & Jeffrey, G. A. *An Introduction to Hydrogen Bonding* (Oxford Univ. Press, 1997).
5. Nagle, J. F. & Morowitz, H. J. Molecular mechanisms for proton transport in membranes. *Proc. Natl Acad. Sci. USA* **75**, 298–302 (1978).
6. Dai, S. et al. Low-barrier hydrogen bonds in enzyme cooperativity. *Nature* **573**, 609–613 (2019).
7. Freier, E., Wolf, S. & Gerwert, K. Proton transfer via a transient linear water-molecule chain in a membrane protein. *Proc. Natl Acad. Sci. USA* **108**, 11435–11439 (2011).
8. Bertalan, É., Lešnik, S., Bren, U. & Bondar, A.-N. Protein-water hydrogen-bond networks of G protein-coupled receptors: graph-based analyses of static structures and molecular dynamics. *J. Struct. Biol.* **212**, 107634 (2020).
9. Shinobu, A. & Agmon, N. Proton wire dynamics in the green fluorescent protein. *J. Chem. Theory Comput.* **13**, 353–369 (2017).
10. Bjerregaard-Andersen, K. et al. A proton wire and water channel revealed in the crystal structure of isatin hydrolase. *J. Biol. Chem.* **289**, 21351–21359 (2014).
11. Cleland, W. W., Frey, P. A. & Gerlt, J. A. The low barrier hydrogen bond in enzymatic catalysis. *J. Biol. Chem.* **273**, 25529–25532 (1998).
12. Oltrogge, L. M. & Boxer, S. G. Short hydrogen bonds and proton delocalization in green fluorescent protein (GFP). *ACS Cent. Sci.* **1**, 148–156 (2015).
13. Wang, L., Fried, S. D., Boxer, S. G. & Markland, T. E. Quantum delocalization of protons in the hydrogen-bond network of an enzyme active site. *Proc. Natl Acad. Sci. USA* **111**, 18454–18459 (2014).
14. Wolf, S., Freier, E. & Gerwert, K. A delocalized proton-binding site within a membrane protein. *Biophysical J.* **107**, 174–184 (2014).
15. Graham, J. D., Buytendyk, A. M., Wang, D., Bowen, K. H. & Collins, K. D. Strong, low-barrier hydrogen bonds may be available to enzymes. *Biochemistry* **53**, 344–349 (2014).
16. Hosur, M. V. et al. Low-barrier hydrogen bonds in proteins. *Crystallogr. Rev.* **19**, 3–50 (2013).
17. Nichols, D. A. et al. Ligand-induced proton transfer and low-barrier hydrogen bond revealed by X-ray crystallography. *J. Am. Chem. Soc.* **137**, 8086–8095 (2015).
18. Tripathi, R., Forbert, H. & Marx, D. Settling the long-standing debate on the proton storage site of the prototype light-driven proton pump bacteriorhodopsin. *J. Phys. Chem. B* **123**, 9598–9608 (2019).
19. Oesterhelt, D. & Stoeckenius, W. Rhodopsin-like protein from the purple membrane of Halobacterium halobium. *Nat. N. Biol.* **233**, 149–152 (1971).

20. Henderson, R. et al. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **213**, 899–929 (1990).
21. Lanyi, J. K. Bacteriorhodopsin. *Annu. Rev. Physiol.* **66**, 665–688 (2004).
22. Chizhov, I. et al. Spectrally silent transitions in the bacteriorhodopsin photocycle. *Biophysical J.* **71**, 2329–2345 (1996).
23. Stoeckenius, W. & Rowen, R. A morphological study of *Halobacterium halobium* and its lysis in media of low salt concentration. *J. Cell Biol.* **34**, 365–393 (1967).
24. Henderson, R. & Unwin, P. N. T. Th ee-dimensional model of purple membrane obtained by electron microscopy. *Nature* **257**, 28–32 (1975).
25. Zaccai, G. & Gilmore, D. J. Areas of hydration in the purple membrane of *Halobacterium halobium*: a neutron diffraction study. *J. Mol. Biol.* **132**, 181–191 (1979).
26. Papadopoulos, G., Dencher, N. A., Zaccai, G. & Bu¨ldt, G. Water molecules and exchangeable hydrogen ions at the active centre of bacteriorhodopsin localized by neutron diffraction: elements of the proton pathway? *J. Mol. Biol.* **214**, 15–19 (1990).
27. Landau, E. M. & Rosenbusch, J. P. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc. Natl Acad. Sci. USA* **93**, 14532–14535 (1996).
28. Pebay-Peyroula, E., Rummel, G., Rosenbusch, J. P. & Landau, E. M. X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science* **277**, 1676–1681 (1997).
29. Inoue, K. et al. A light-driven sodium ion pump in marine bacteria. *Nat. Commun.* **4**, 1678 (2013).
30. Nagel, G. et al. Channelrhodopsin-1: a light-gated proton channel in green algae. *Science* **296**, 2395–2398 (2002).
31. Nagel, G. et al. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc. Natl Acad. Sci. USA* **100**, 13940–13945 (2003).
32. Shevchenko, V. et al. Inward H^+ pump xenorhodopsin: mechanism and alternative optogenetic approach. *Sci. Adv.* **3**, e1603187 (2017).
33. Inoue, K. et al. A natural light-driven inward proton pump. *Nat. Commun.* **7**, 13415 (2016).
34. Pushkarev, A. et al. A distinct abundant group of microbial rhodopsins discovered using functional metagenomics. *Nature* **558**, 595–599 (2018).
35. Yutin, N. & Koonin, E. V. Proteorhodopsin genes in giant viruses. *Biol. Direct* **7**, 34 (2012).
36. Hoffmann, A., Hildebrandt, V., Heberle, J. & Buldt, G. Photoactive mitochondria: in vivo transfer of a light-driven proton pump into the inner mitochondrial membrane of *Schizosaccharomyces pombe*. *Proc. Natl Acad. Sci. USA* **91**, 9367–9371 (1994).
37. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. & Lanyi, J. K. Structure of bacteriorhodopsin at 1.55 Å resolution. *J. Mol. Biol.* **291**, 899–911 (1999).
38. Edman, K. et al. High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature* **401**, 822–826 (1999).
39. Royant, A. et al. Helix deformation is coupled to vectorial proton transport in the photocycle of bacteriorhodopsin. *Nature* **406**, 645–648 (2000).
40. Matsui, Y. et al. Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin. *J. Mol. Biol.* **324**, 469–481 (2002).
41. Kouyama, T., Nishikawa, T., Tokuhisa, T. & Okumura, H. Crystal structure of the L intermediate of bacteriorhodopsin: evidence for vertical translocation of a water molecule during the proton pumping cycle. *J. Mol. Biol.* **335**, 531–546 (2004).
42. Takeda, K. et al. Crystal structure of the M intermediate of bacteriorhodopsin: allosteric structural changes mediated by sliding movement of a transmembrane helix. *J. Mol. Biol.* **341**, 1023–1037 (2004).
43. Sass, H. J. et al. Structural alterations for proton translocation in the M state of wild-type bacteriorhodopsin. *Nature* **406**, 649–653 (2000).
44. Yamamoto, M., Hayakawa, N., Murakami, M. & Kouyama, T. Crystal structures of different substates of bacteriorhodopsin's m intermediate at various pH levels. *J. Mol. Biol.* **393**, 559–573 (2009).
45. Wickstrand, C., Dods, R., Royant, A. & Neutze, R. Bacteriorhodopsin: would the real structural intermediates please stand up? *Biochimica et. Biophysica Acta (BBA) - Gen. Subj.* **1850**, 536–553 (2015).
46. Gerwert, K., Freier, E. & Wolf, S. The role of protein-bound water molecules in microbial rhodopsins. *Biochimica et. Biophysica Acta (BBA) - Bioenerg.* **1837**, 606–613 (2014).
47. Maeda, A., Gennis, R. B., Balashov, S. P. & Ebrey, T. G. Relocation of water molecules between the schiff base and the Thr4 −Asp96 region during light-driven unidirectional proton transport by bacteriorhodopsin: an FTIR study of the N intermediate. *Biochemistry* **44**, 5960–5968 (2005).
48. Borshchevskiy, V., Efremov, R., Moiseeva, E., Büldt, G. & Gordeliy, V. Overcoming merohedral twinning in crystals of bacteriorhodopsin grown in lipidic mesophase. *Acta Cryst. D.* **66**, 26–32 (2010).
49. Borshchevskiy, V. I., Round, E. S., Popov, A. N., Büldt, G. & Gordeliy, V. I. X-ray-radiation-induced changes in bacteriorhodopsin structure. *J. Mol. Biol.* **409**, 813–825 (2011).
50. Nango, E. et al. A three-dimensional movie of structural changes in bacteriorhodopsin. *Science* **354**, 1552–1557 (2016).
51. Nogly, P. et al. Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science* **361**, 845–851 (2018).
52. Nass Kovacs, G. et al. Th ee-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin. *Nat. Commun.* **10**, 3177 (2019).
53. Kachalova, G. S., Popov, A. N. & Bartunik, H. D. A steric mechanism for inhibition of CO binding to heme proteins. *Science* **284**, 473–476 (1999).
54. Hasegawa, N., Jonotsuka, H., Miki, K. & Takeda, K. X-ray structure analysis of bacteriorhodopsin at 1.3 Å resolution. *Sci. Rep.* **8**, 13123 (2018).
55. Gordeliy, V. I. et al. Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex. *Nature* **419**, 484–487 (2002).
56. Gordeliy, V. I., Schlesinger, R., Efremov, R., Büldt, G. & Heberle, J. in *Membrane Protein Protocols: Expression, Purification, and Characterization* (ed. Selinsky, B. S.) 305–316 (Humana Press, 2003); https://doi.org/10.1385/1-59259-400-X:305
57. Borshchevskiy, V. et al. Low-dose X-ray radiation induces structural alterations in proteins. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **70**, 2675–2685 (2014).
58. Moukhametzianov, R. et al. Development of the signal in sensory rhodopsin and its transfer to the cognate transducer. *Nature* **440**, 115–119 (2006).
59. Kovalev, K. et al. Molecular mechanism of light-driven sodium pumping. *Nat. Commun.* **11**, 2137 (2020).
60. Gushchin, I. et al. Active state of sensory rhodopsin II: structural determinants for signal transfer and proton pumping. *J. Mol. Biol.* **412**, 591–600 (2011).
61. Frey, P. A. in *Encyclopedia of Biological Chemistry* 594–598 (Elsevier, 2004); https://doi.org/10.1016/B0-12-443710-9/00367-7
62. Garczarek, F. & Gerwert, K. Functional waters in intraprotein proton transfer monitored by FTIR difference spectroscopy. *Nature* **439**, 109–112 (2006).
63. Rammelsberg, R., Hessling, B., Chorongiewski, H. & Gerwert, K. Molecular reaction mechanisms of proteins monitored by nanosecond step-scan FT-IR difference spectroscopy. *Appl. Spectrosc.*, **51**, 558–562 (1997).
64. Zscherp, C., Schlesinger, R., Tittor, J., Oesterhelt, D. & Heberle, J. In situ determination of transient pKa changes of internal amino acids of bacteriorhodopsin by using time-resolved attenuated total refl ction Fourier-transform infrared spectroscopy. *Proc. Natl Acad. Sci. USA* **96**, 5498–5503 (1999).
65. Garczarek, F., Brown, L. S., Lanyi, J. K. & Gerwert, K. Proton binding within a membrane protein by a protonated water cluster. *Proc. Natl Acad. Sci. USA* **102**, 3633–3638 (2005).
66. Goyal, P. et al. Proton storage site in bacteriorhodopsin: new insights from quantum mechanics/molecular mechanics simulations of microscopic pK_a and infrared spectra. *J. Am. Chem. Soc.* **133**, 14981–14997 (2011).
67. Bada Juarez, J. F. et al. Structures of the archaerhodopsin-3 transporter reveal that disordering of internal water networks underpins receptor sensitization. *Nat. Commun.* **12**, 629 (2021).
68. Palczewski, K. et al. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745 (2000).
69. Ballesteros, J. A. et al. Activation of the β2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6*. *J. Biol. Chem.* **276**, 29171–29177 (2001).
70. Schneider, E. H., Schnell, D., Strasser, A., Dove, S. & Seifert, R. Impact of the DRY motif and the missing 'ionic lock' on constitutive activity and G-protein coupling of the human histamine H4 receptor. *J. Pharm. Exp. Ther.* **333**, 382–392 (2010).
71. Lorenz-Fonfria, V. A., Saita, M., Lazarova, T., Schlesinger, R. & Heberle, J. pH-sensitive vibrational probe reveals a cytoplasmic protonated cluster in bacteriorhodopsin. *Proc. Natl. Acad. Sci. USA* **114**, E10909–E10918 (2017).
72. Nakamura, A. et al. 'Newton's cradle' proton relay with amide–imidic acid tautomerization in inverting cellulase visualized by neutron crystallography. *Sci. Adv.* **1**, e1500263 (2015).
73. Heberle, J., Riesle, J., Thi demann, G., Oesterhelt, D. & Dencher, N. A. Proton migration along the membrane surface and retarded surface to bulk transfer. *Nature* **370**, 379–382 (1994).
74. Maeda, A., Sasaki, J., Shichida, Y. & Yoshizawa, T. Water structural changes in the bacteriorhodopsin photocycle: analysis by Fourier transform infrared spectroscopy. *Biochemistry* **31**, 462–467 (1992).
75. Zaccai, G. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science* **288**, 1604–1607 (2000).
76. Subramaniam, S. & Henderson, R. Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature* **406**, 653–657 (2000).
77. Weinert, T. et al. Proton uptake mechanism in bacteriorhodopsin captured by serial synchrotron crystallography. *Science* **365**, 61–65 (2019).
78. Lanyi, J. K. & Váró, G. The photocycles of bacteriorhodopsin. *Isr. J. Chem.* **35**, 365–385 (1995).
79. Lanyi, J. K. & Schobert, B. Crystallographic structure of the retinal and the protein after deprotonation of the schiff base: the switch in the bacteriorhodopsin photocycle. *J. Mol. Biol.* **321**, 727–737 (2002).

80. Ho, B. K. & Gruswitz, F. HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct. Biol.* **8**, 49 (2008).
81. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).
82. Liebschner, D. et al. Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. D. Struct. Biol.* **73**, 148–157 (2017).

Methods

Protein crystallization. The expression and purifi ation of BR are described in detail in ref. [56]. In brief, purple membranes were isolated from *H. salinarum* strain S9 and BR was isolated from purple membranes by solubilization in CYMAL-5 (5-cyclohexyl-1-pentyl-d-maltoside) detergent (Affimetrix). BR crystals were grown in a lipidic cubic phase of monoolein (Nu-Chek Prep). In accordance with the described procedure, in the case of BR crystallization trials were set up in PCR tubes containing 15 µl of a mesophase of monoolein and the protein in CYMAL-5. A mixture of dry salts (Na_2HPO_4 (5%) and KH_2PO_4 (95%)) was used as a precipitant. Crystallization probes were incubated at 22 °C for 1–2 months. Flat hexagonally shaped plates of 250–600 µm along the plane and 10–70 µm in thickness were obtained. To separate crystals from the lipidic mesophase, 3 M sodium phosphate buffer pH 5.6 with 0.1% *n*-octyl-β-d-glucopyranoside was used. Crystals of BR were harvested using micromounts (Mitegen), flash-frozen and stored in liquid nitrogen.

Intermediate states trapping procedures. All X-ray data for BR were collected at ESRF crystallographic beamlines (ID14-1 and ID14-2). For the ground-state data collection, the crystal was light adapted and cryogenically cooled to 100 K. The fraction of radiationally damaged protein was estimated as described in ref. [49] and did not exceed 5%. All intermediate states were trapped immediately at the beamlines with crystals installed at the goniometer. Crystallographic data for the K state data were collected similar to ref. [40]. In brief, three sets of partial diffraction data from a single crystal covering 11° of the reciprocal space were collected at 100 K in sequence under alternating illumination conditions: I, red light (676 nm, 0.7 mW); II, green light (530 nm, 0.3 mW) and III, red light (676 nm, 0.7 mW). Such measurements were repeated for six consequent orientations of the crystal, and data collected under the illumination conditions I, II and III were merged separately to provide three complete datasets from which experimental difference maps $F_I - F_{III}$, $F_{II} - F_{III}$ and $F_{II} - F_{III}$ were calculated. Images taken in the initial X-ray exposure before saturation of the orange species were excluded[57]. $F_I - F_{III}$ shows no difference confirming no detectable radiation damage or orange species formation. Meanwhile, $F_{II} - F_{III/I}$ contain differential densities that we associate with the K state reversibly formed under green laser illumination. The cryogenic trapping procedure was independently validated at the off-line microspectrophotometer[83]. The spectra were drawn using Microsoft Excel (v.16.58). Illumination of the BR crystal with a green laser (532 nm, 10 mW mm^{-2}) at 100 K confirmed the formation of the K state.

For the L state trapping, the cryogenically cooled BR crystal was gradually warmed up to 160 K and illuminated with a red laser (676 nm, 0.7 mW) for 5 min; after that, the crystal was gradually cooled down back to 100 K and the X-ray dataset was collected. The fraction of radiationally damaged protein was estimated as described in ref. [49] and did not exceed 5%. The cryogenic trapping procedure was independently validated at the off-line microspectrophotometer[83]. Illumination of the BR crystal with a red laser (633 nm, 3 mW mm^{-2}) at 100 K confirms the predominant formation of the L state with the minor fraction of the M state (<10%).

The M state of BR was trapped in a similar way to ref. [43]. The cryogenically cooled crystal was heated to room temperature by blocking the nitrogen cold stream for 3 s while illuminating with a green light (33 mW, 514 nm), then cooled to 100 K by unblocking the cold stream while the green light was still on. One second after recooling started, the illumination was turned off.

Model building and refinement. The X-ray diffraction data were collected using ProDC and MxCube software. Integration and scaling of the BR data were performed in XDS (BUILT = 20210205)[84]. $CC_{1/2}$ on the outer shell was used as the main criterion of resolution cutoff as recommended[85]. The twinning ratio was determined using Yeates statistics and the DETWIN routine of CCP4i program suite (v.7.1.018)[86,87] as described previously[48]. The BR ground-state model with PDB ID 1C3W (ref. [37]) was converted to a polyalanine model. Rigid-body refinement and automated macromolecular model building with subsequent refinement were performed in ARP/wARP (v.8.0)[88], REFMAC5 (v.5.8.0267)[89], PHENIX (v.1.19.2-4158)[90] and Coot (v.0.9.6)[91]. Experimental difference Fourier Q-weighted electron-density maps[92] between intermediate and ground states were calculated in CNS (v.1.3)[93] with phases from the ground-state model. Initial models for intermediate states were built in accordance with difference maps and iteratively refined manually in Coot and automatically in PHENIX. The structural figures were prepared using PyMoL (v.1.8.4.0).

Identification of LBHBs. The difference electron-density maps do not show hydrogen atoms in the true-atomic-resolution structures of BR obtained in the present work. Therefore, for the tracing of LBHBs we analyzed the lengths of H-bonds in all states of BR. For accurate analysis, the coordinate errors of the relative positions of the atoms were calculated as proposed in ref. [94]. In brief, the position (r) error of atom X was estimated as $\sigma(r, B_X) = (B_X/B_{avg})^{1/2} 3^{1/2} (N_i/n_{obs})^{1/2} C \cdot 1/3 R_{free} d_{min}$, where B_X and B_{avg} are B factors of atom X and mean value for all atoms in the structure, respectively, N_i is total number of atoms in structure, n_{obs} is a number of reflections used for refinement and C is the completeness of the data to a resolution limit of d_{min}. The distance error for atoms X and Y was then estimated as $\sigma_d = (\sigma_X^2 + \sigma_Y^2)^{1/2}$. For the true-atomic-resolution structures of BR in the ground, L and M states, the σ_d values were at the level of 0.1 Å. In the structure of the K state of BR at 1.53 Å, the distance error was higher (roughly 0.2 Å).

The H-bonds with an O–O distance of >2.6 Å or N–O distance >2.7 Å are considered normal H-bonds, while the H-bonds with the O–O or N–O distances of 2.3–2.6 Å are considered to be short strong H-bonds (LBHBs)[4,61]. Taking into account the estimated experimental error of the H-bond length determination of 0.1 Å, we assigned only the H-bonds with O–O and N–O distances of 2.3–2.5 Å to LBHBs. It should be noted that among all short H-bonds identified in the presented structures, the lengths of most were in the range of 2.4–2.5 Å, which allowed us to unequivocally assign these bonds to the short strong H-bond group (LBHBs). There is only one extremely short (2.3 Å) bond between E194 and E204 in the L state of BR, which, taking the experimental error of 0.1 Å into account, could be also interpreted as a single-well H-bond (the length of such bonds is less than 2.3 Å). However, this assumption is very speculative and more experiments should be performed to validate the nature of this H-bond.

Structure-based bioinformatic analysis. A collection of representative microbial rhodopsin sequences was aligned with hmmalign and used to construct a profile hidden Markov model using hmmbuild (v.hmmer-3.3.2)[95]. The dataset of BRs was constructed by searching the UniProt database using jackhmmer (ten iterations) with the obtained hidden Markov model profile and consecutive selection of archaeal genes with characteristic motif RDTDDK (82, 85, 89, 96, 212 and 216 positions according to BR numbering). In total, 270 BRs were selected and aligned using the CLUSTALO online server[96]. The phylogenetic tree was generated using the iTol v.6 online server[97]. The analyzed regions were chosen based on the X-ray structures of BR presented in this work. Extended Data Fig. 10 preparations were done using Python (v.3.6.9), and in-house Jupyter (v.5.2.2) and Wolfram Mathematica (v.11) Notebooks.

Data availability

Atomic models built using X-ray crystallography data have been deposited in the Research Collaboratory for Structural Bioinformatics PDB with the codes 7Z09 (for the ground state of BR at 1.05 Å), 7Z0A (for the ground state of BR at 1.22 Å), 7Z0C (for the K state), 7Z0D (for the L state) and 7Z0E (for the M state). The UniProt gene database (uniport.org) was used for bioinformatic analysis of the BR clade of microbial rhodopsins.

References

83. Efremov, R., Gordeliy, V. I., Heberle, J. & Büldt, G. Time-resolved microspectroscopy on a single crystal of bacteriorhodopsin reveals lattice-induced differences in the photocycle kinetics. *Biophysical J.* **91**, 1441–1451 (2006).
84. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
85. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
86. Yeates, T. O. in *Methods in Enzymology* Vol. 276, 344–358 (Academic Press, 1997).
87. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Cryst. D.* **67**, 235–242 (2011).
88. Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. ARP/wARP and molecular replacement. *Acta Cryst. D.* **57**, 1445–1450 (2001).
89. Murshudov, G. N. et al. REFMAC5 for the refi ement of macromolecular crystal structures. *Acta Crystallogr. Sect. D., Biol. Crystallogr.* **67**, 355–367 (2011).
90. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
91. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
92. Bourgeois, D. New processing tools for weak and/or spatially overlapped macromolecular diffraction patterns. *Acta Cryst. D.* **55**, 1733–1741 (1999).
93. Brünger, A. T. et al. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Cryst. D.* **54**, 905–921 (1998).
94. Cruickshank, D. W. J. Remarks about protein structure precision. *Acta Crystallogr. D. Biol. Crystallogr.* **55**, 583–601 (1999).
95. Wheeler, T. J. & Eddy, S. R. nhmmer: DNA homology search with profile HMMs. *Bioinformatics* **29**, 2487–2489 (2013).
96. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
97. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).

98. von Stetten, D. et al. In crystallo optical spectroscopy (icOS) as a complementary tool on the macromolecular crystallography beamlines of the ESRF. *Acta Crystallogr. D. Biol. Crystallogr.* **71**, 15–26 (2015).

Acknowledgements

We acknowledge the Structural Biology Group of the ESRF for granting access to the synchrotron beamlines. We are deeply thankful to A. Round for his help with manuscript preparation. This work was supported by the common program of Agence Nationale de la Recherche (ANR), France and Deutsche Forschungsgemeinschaft, Germany (grant no. ANR-15-CE11-0029-02) and by funding from Frankfurt: Cluster of Excellence Frankfurt Macromolecular Complexes (to E.B.) by the Max Planck Society (to E.B.) and by the Commissariat à l'Energie Atomique et aux Energies Alternatives (Institut de Biologie Structurale)–Helmholtz-Gemeinschaft Deutscher Forschungszentren (Forschungszentrum Jülich) Special Terms and Conditions 5.1 specific agreement. V.G. greatly acknowledges his HGF Professorship. V.B. acknowledges DAAD Young Talents Programme Line A. This work used the *ic*OS and HTX platforms of the Grenoble Instruct-ERIC center (ISBG; UMS3518 CNRS-CEA-UJF-EMBL) within the Grenoble Partnership for Structural Biology (PSB)[95]. Platform access was supported by FRISBI (grant no. ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (grant no. ANR-17-EURE-0003). Protein crystallization was supported by the Russian Foundation for Basic Research (RFBR) according to the research project no. 18-02-40020. Data collections for the K intermediate state structures were supported by RFBR (project no. 19-29-12022). Data collections for the L intermediate state structures were supported by the Russian Ministry of Science and Higher Education (grant no. 075-15-2021-1354). Data collection of ground state was supported by RFBR-CNRS (project no. 19-52-15017). Data treatment and structure solution were supported by the Russian Science Foundation (RSF) (project no. 21-64-00018). Bioinformatic analysis was supported by RSF-Helmholtz (grant no. 19-44-06302). The work is supported by Ministry of Science and Higher Education of the Russian Federation (project no. 075-00958-21-05/730000F. 99.1.BV10AA00006 to A.R. for data analysis, and project 075-00337-20-03/FSMG-2020-0003 to V.B. for X-ray data treatment).

Author contributions

C.B. and D.B. expressed and purified the protein. T.B. supervised the expression and purification. V.G., E.R. and V.B. crystallized the protein. R.A. helped with the crystallization. R.E. and V.B. collected absorption spectra from crystals and performed cryo-trapping of the intermediates. V.G. supervised the absorption spectra collection. V.B. and E.R. collected the diffraction data. V.B. and K.K. solved the structures. K.K. performed final refinement of the structures with the help of V.B. and G. Bourenkov. I.G. helped with structure analysis. A.R., V.B., E.B., M.E., I.C., D.W., A.K. and G.B. helped with data analysis. A.A. performed bioinformatic analysis of the archaeal outward proton-pumping rhodopsins. V.G. developed crystallization approaches allowing growth of high-quality crystals. V.G. designed and supervised the project with contribution of G. Büldt. and analyzed major structure-function relationships. K.K. and V.G. analyzed the results and prepared the manuscript with the important contribution of V.B. and with input from all the other authors.

Competing interests

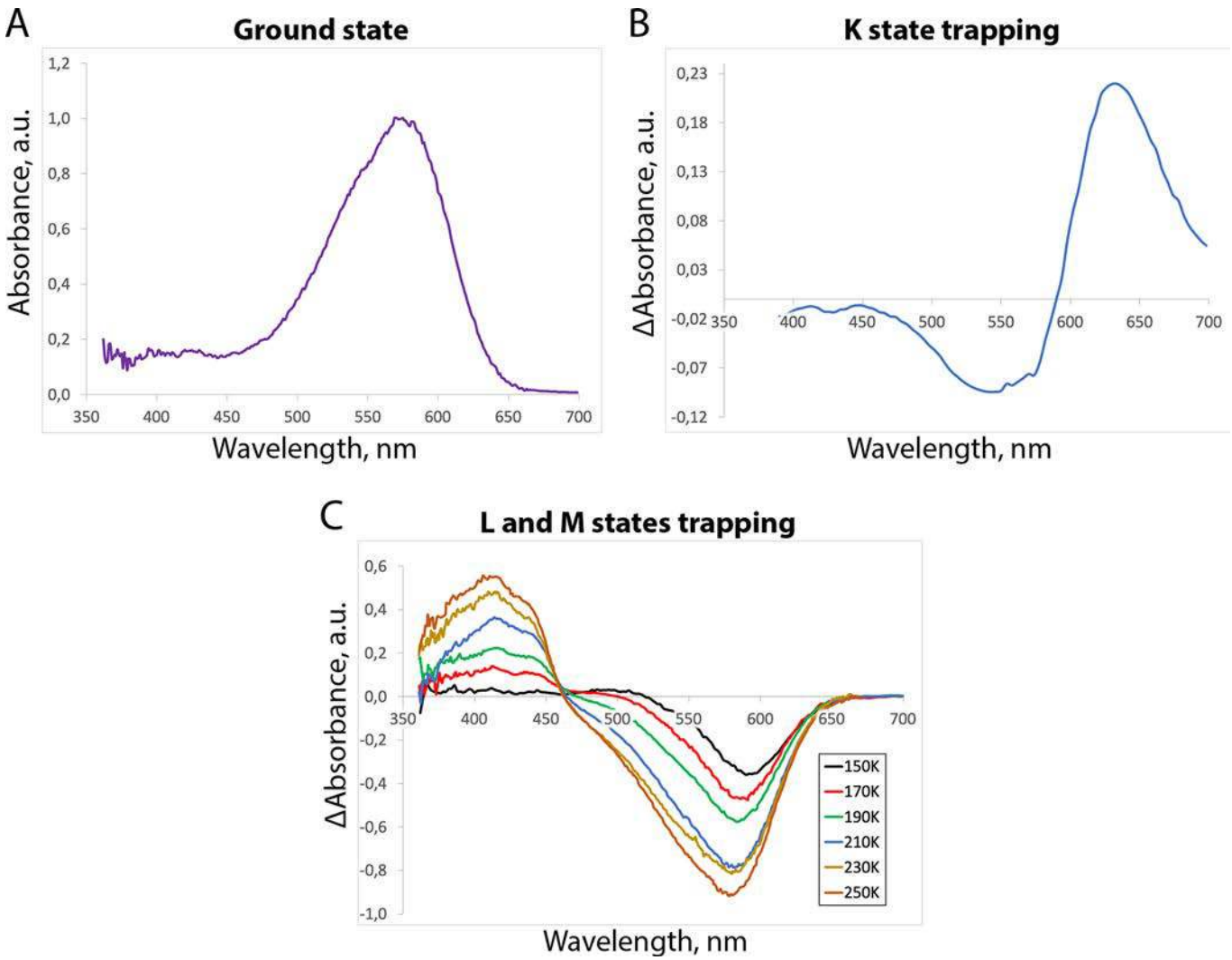
The authors declare no competing interests.

Additional information

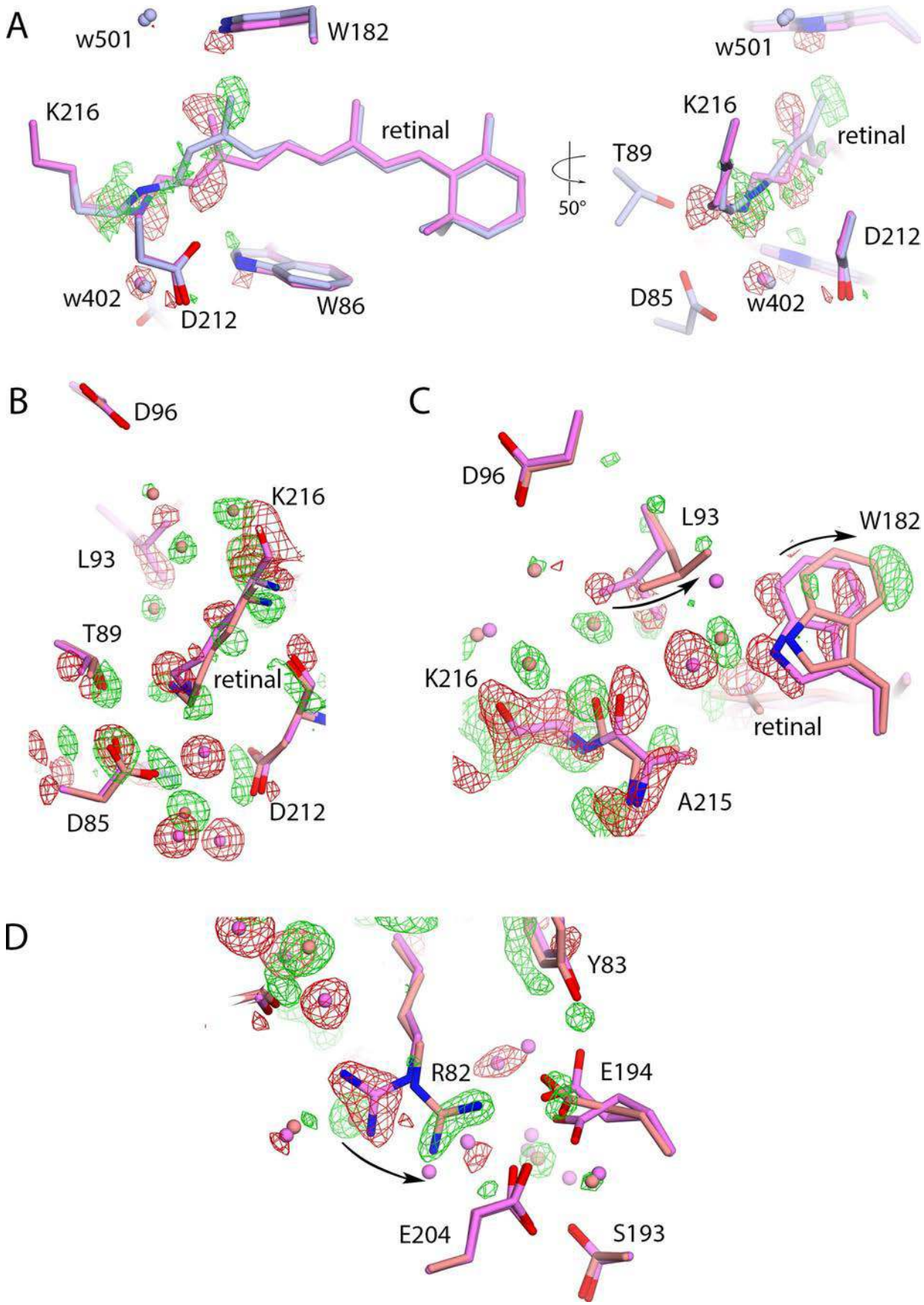
Extended data are available for this paper at https://doi.org/10.1038/s41594-022-00762-2.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-022-00762-2.

Correspondence and requests for materials should be addressed to Valentin Gordeliy.

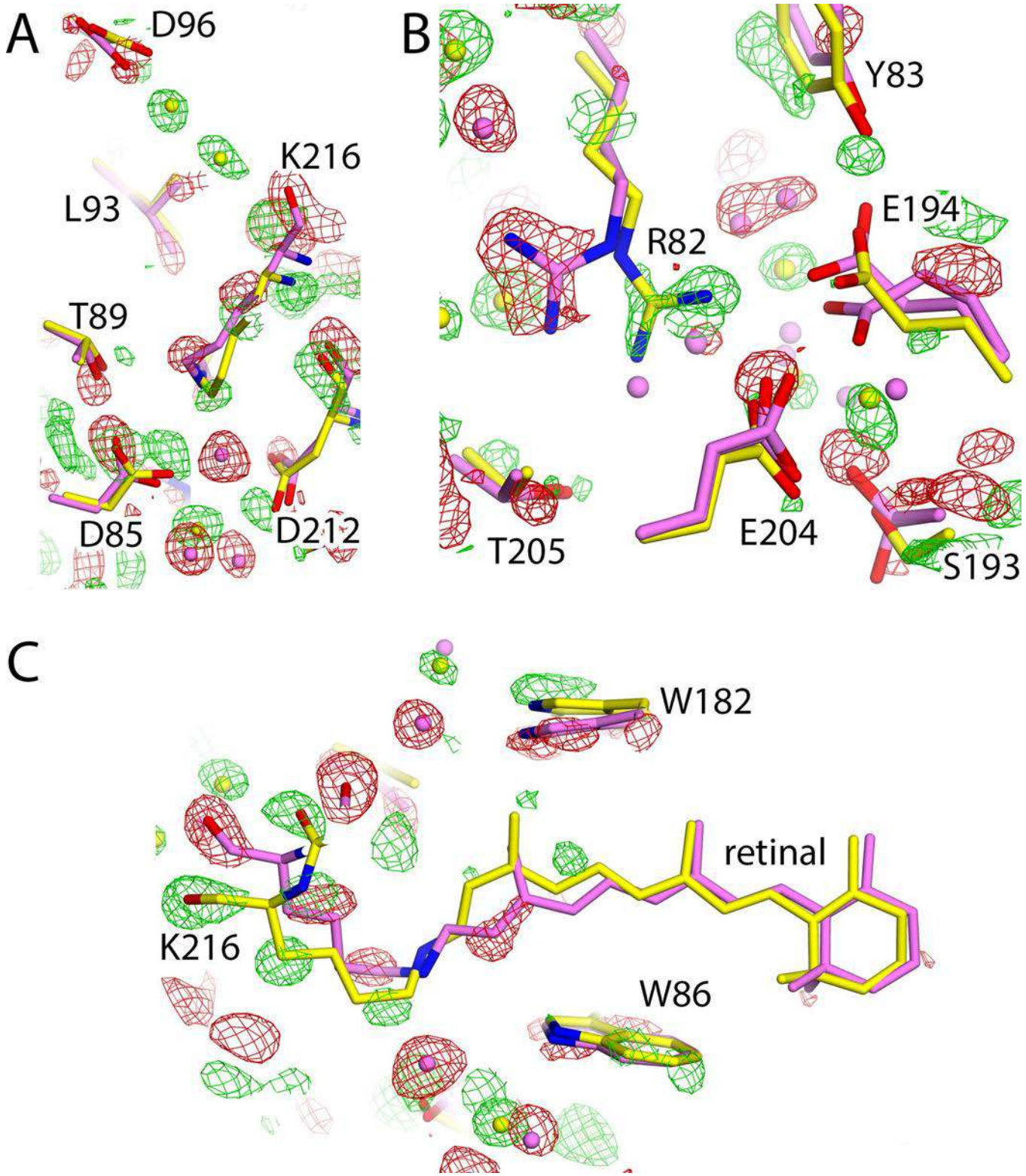


Extended Data Fig. 1 | Spectroscopy of BR crystals and validation of the cryotrapped intermediates. A. Spectrum of the ground state of BR in a crystal at 100K. **B**. Difference absorption spectrum (final minus ground) of after cryotrapping of the K state in BR crystals at 100K. **C**. Difference absorption spectra (final minus ground) after the cryotrapping of the L and M states in BR crystals at 150 (black), 170 (red), 190 (green), 210 (blue), 230 (gold), and 250 K (orange).

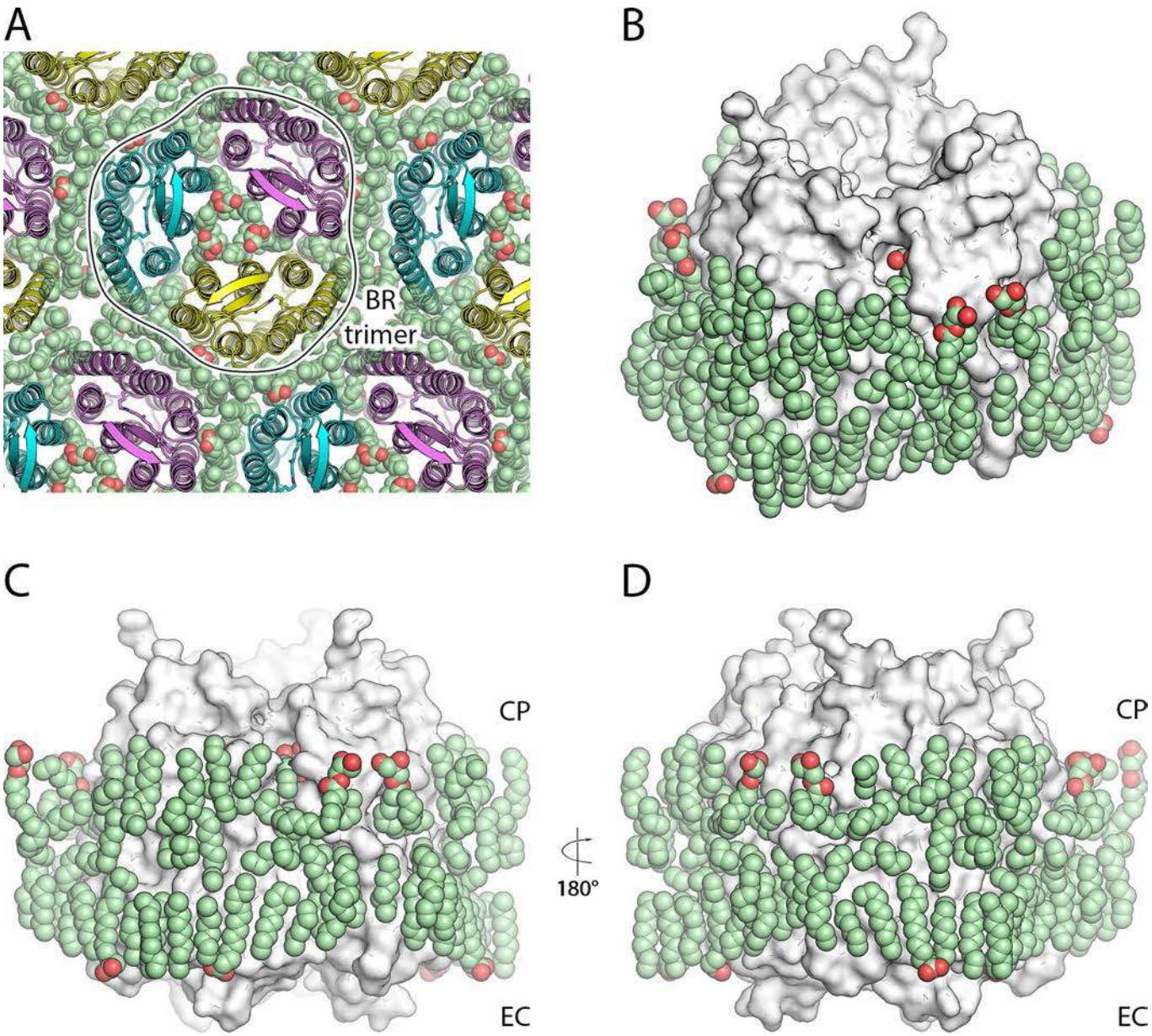


Extended Data Fig. 2 | See next page for caption.

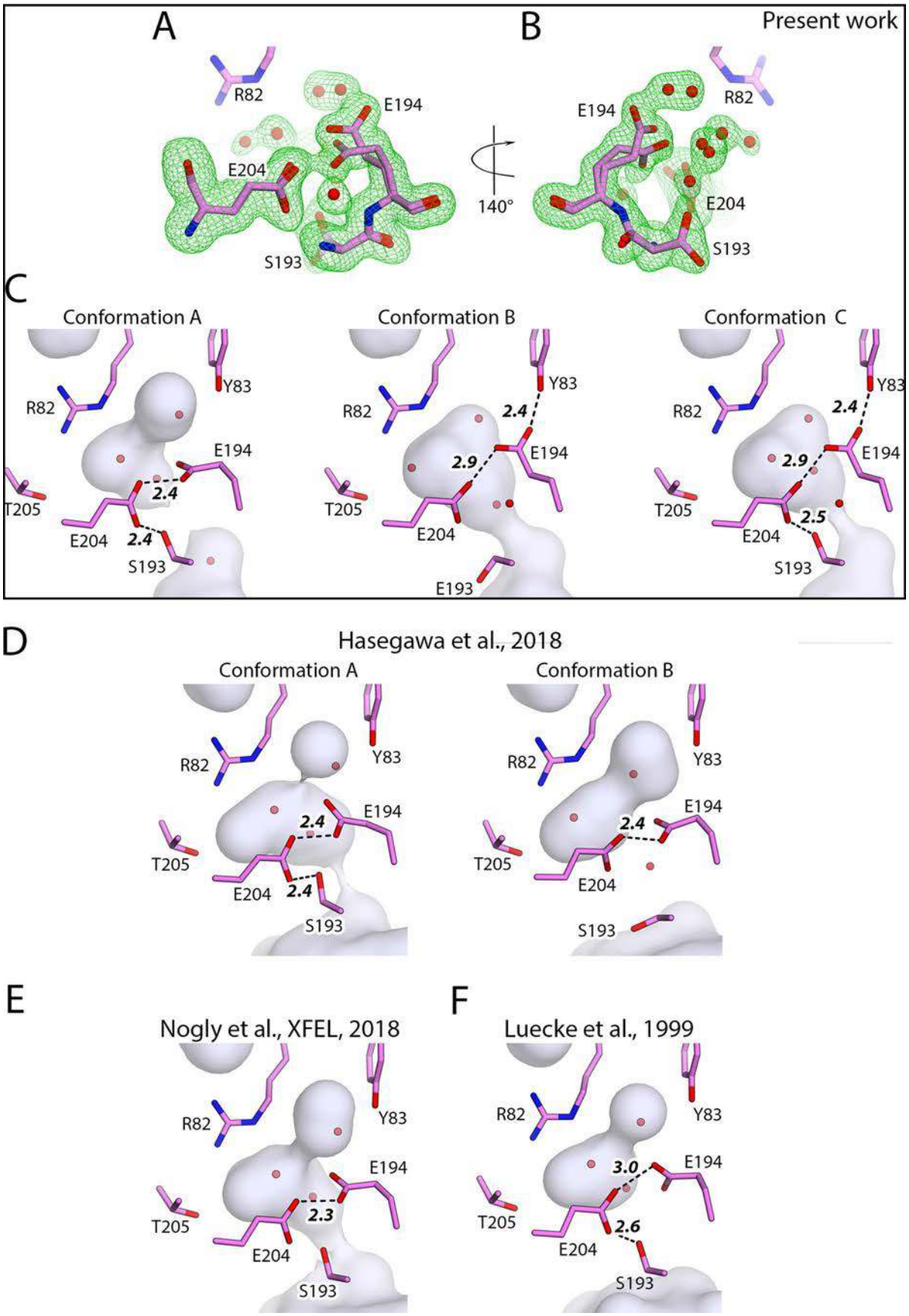
Extended Data Fig. 2 | Examples of the difference F_{oINT}-F_{oGR} electron density maps in BR. A. Difference F_{oK}-F_{oGR} electron density maps built around the retinal cofactor and D212, W86, and W182 residues indicating structural rearrangements in the K state of BR. The maps are contoured at the level of 5σ. Views from two sides are shown. **B**,**C**. Difference F_{oL}-F_{oGR} electron density maps built around the central part of BR indicating structural rearrangements in the L state of BR. The maps are contoured at the level of 3σ. **D**. Difference F_{oL}-F_{oGR} electron density maps built around the R82 residue and proton release group of BR indicating structural rearrangements in the L state of BR. The maps are contoured at the level of 3σ.



Extended Data Fig. 3 | Examples of the difference F_{oM}-F_{oGR} electron density maps in BR. A. Maps built around the central part of BR indicating structural rearrangements in the M state of BR. The maps are contoured at the level of 3σ. **B**. Maps built around the R82 residue and PRG of BR indicating structural rearrangements in the M state of BR. The maps are contoured at the level of 3σ. **C**. Maps built around retinal cofactor and W182 and W86 residues indicating structural rearrangements in retinal binding pocket in the M state of BR. The maps are contoured at the level of 3σ.

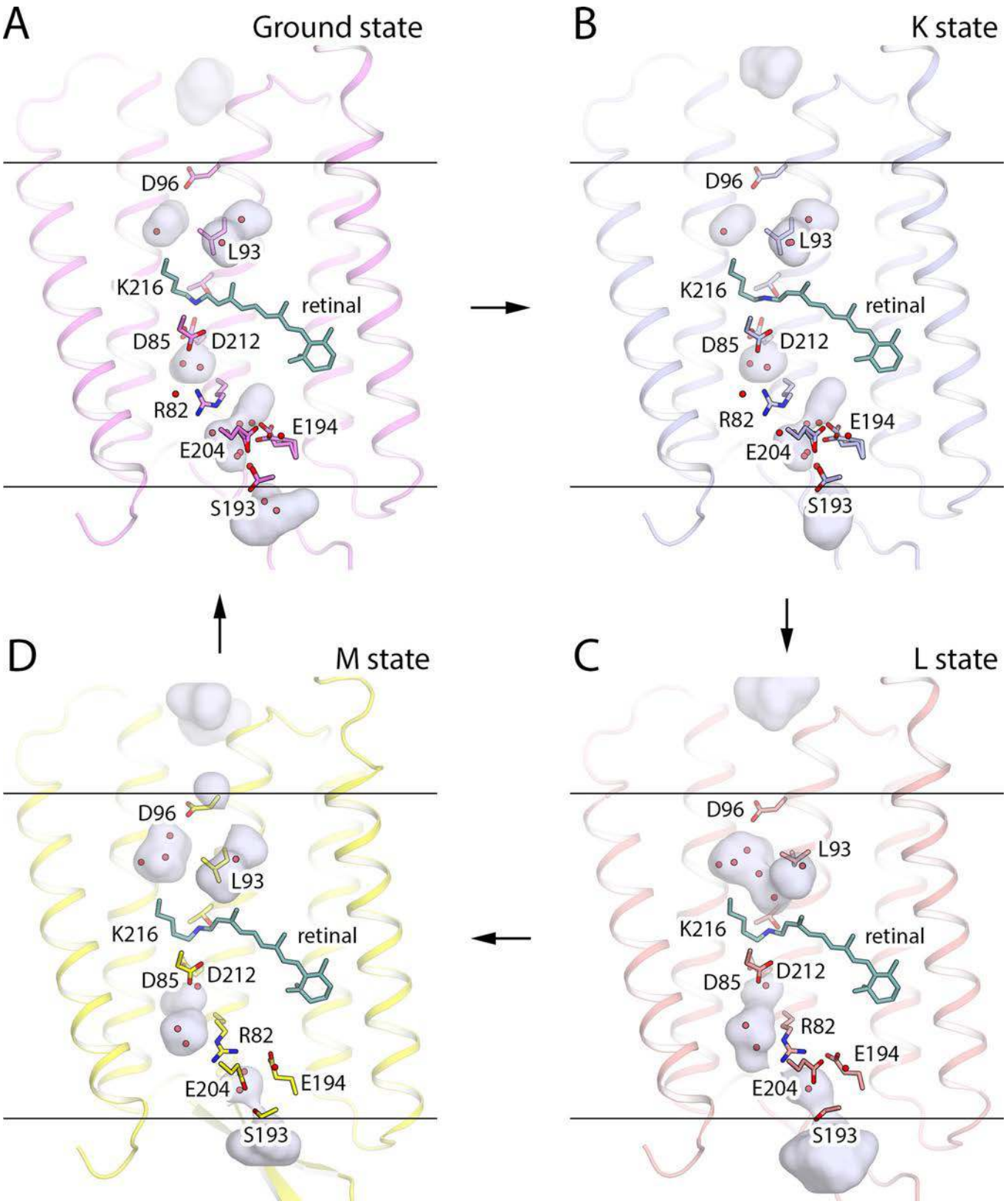


Extended Data Fig. 4 | BR trimer and lipid molecules resolved in the structure. A. View on the crystal monolayer from the extracellular side of BR. The trimer of BR (cartoon representation) is contoured for clarity. **B**. Overall view of one BR trimer (white surface) surrounded by the ring of lipid molecules (green spheres). **C**,**D**. Side views of the BR trimer (white surface) surrounded by the ring of lipid molecules (green spheres).

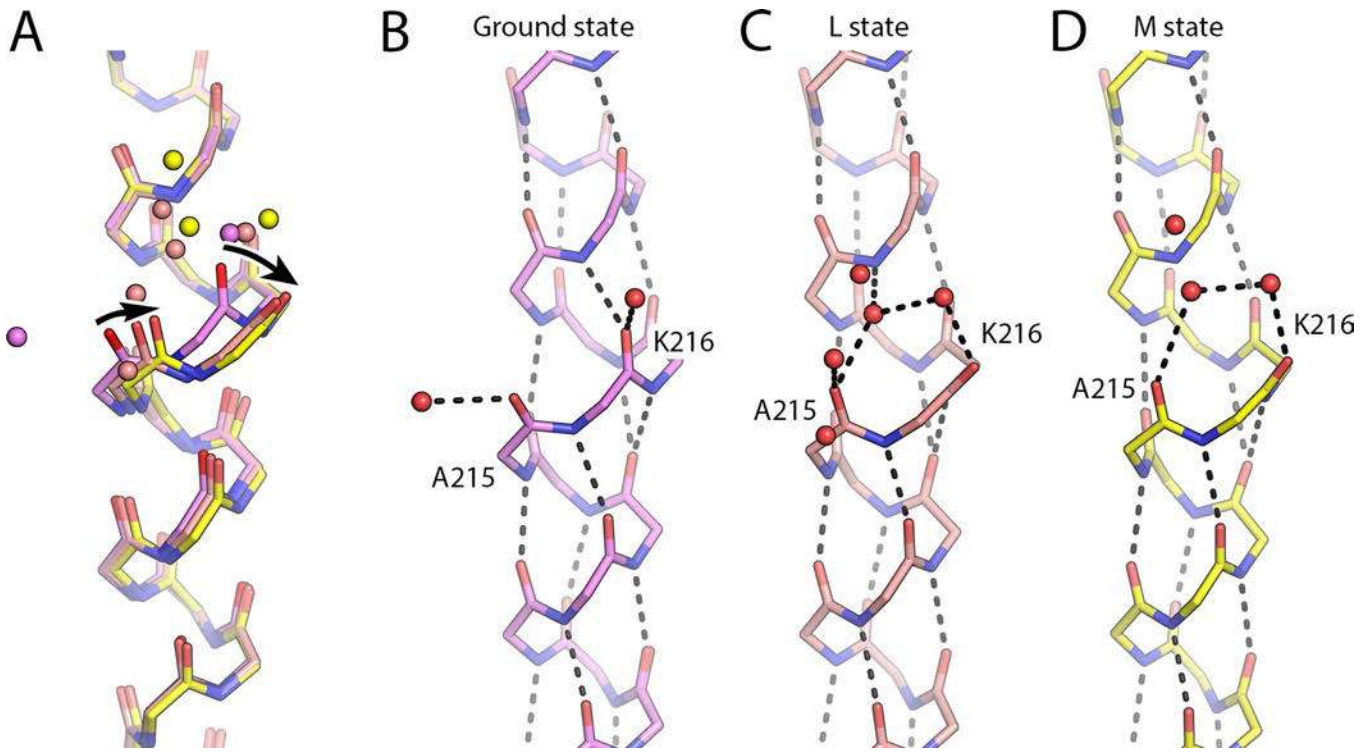


Extended Data Fig. 5 | See next page for caption.

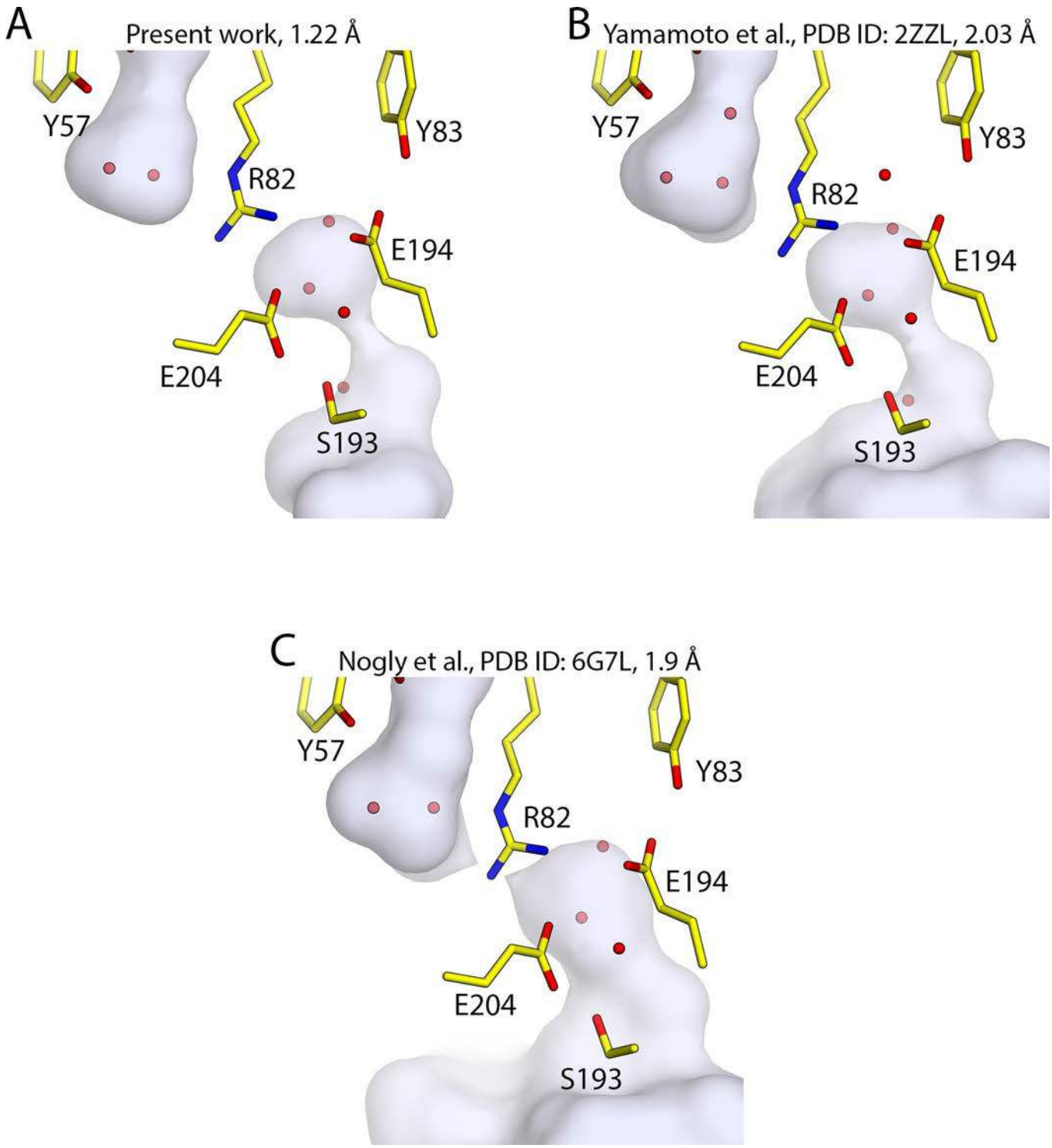
Extended Data Fig. 5 | Structure of the proton release group of BR. A,**B**. Different side views of the PRG in the triple conformation obtained in the present work. The polder maps are contoured at the level of 3σ. **C**. Three conformations coexisting in the model from present work. **D**. Two conformations coexisting in the model from Hasegawa et al.[54] (PDB ID: 5ZIL). **E**. Model of BR obtained using serial femtosecond crystallography (SFX) at XFEL at room temperature from Nogly et al.[51] (PDB ID: 6G7H). **F**. Model of BR from Luecke et al.[37] (PDB ID: 1C3W).



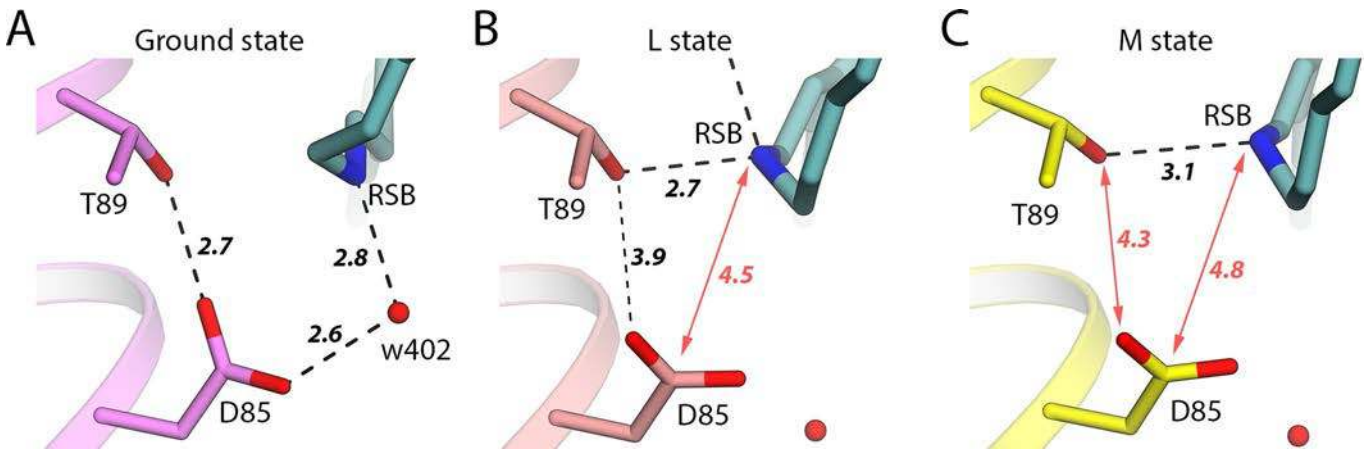
Extended Data Fig. 6 | Overall BR structure and cavities evolution. A. Side view of the ground state (violet). **B**. Side view of the K state (blue). **C**. Side view of the L state (salmon). **D**. Side view of the M state (yellow) of the photocycle. Retinal cofactor is colored teal. Cavities were calculated using HOLLOW and are shown with grey-blue surfaces. Helices F and G are hidden for clarity. Hydrophobic/hydrophilic membrane core boundaries were calculated using PPM server and are shown with black horizontal lines.



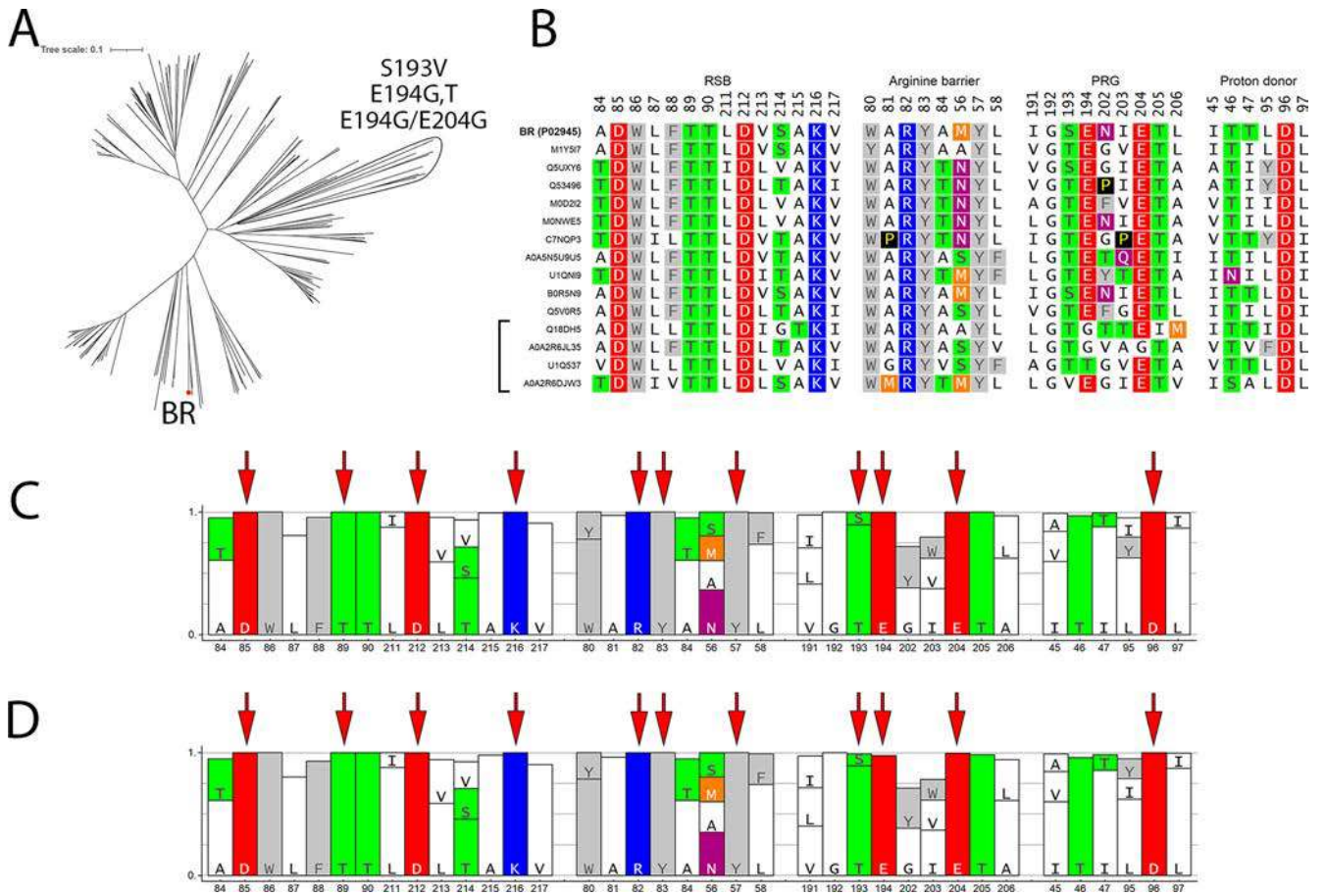
Extended Data Fig. 7 | Helix G of BR in ground, L, and M states. A. Alignment of the backbones of the central part of helix G in the ground (violet), L (salmon), and M (yellow) states. Black arrows indicated most critical rearrangements in the backbones of residues A215 and K216 during photocycle. Water molecules are shown with spheres with the color corresponding to each state. **B**. Backbone of helix G in the ground state of BR. **C**. Backbone of helix G in the L state of BR. **D**. Backbone of helix G in the M state of BR. H-bonds are shown with dashed black lines. Water molecules in panels B-C are shown with red spheres.



Extended Data Fig. 8 | Comparison of the proton release group regions of BR in different structures of the M state. A. The PRG of BR in the M state obtain in present work. **B**. The most similar organization of the PRG of BR found in the M state obtained with cryotrapping procedure[44] (PDB ID: 2ZZL). **C**. The PRG organization in the highest-resolution structure of BR in the M state obtained using time-resolved serial femtosecond crystallography at X-ray Free Electron Laser[51] (PDB ID: 6G7L). Cavities were calculated using HOLLOW and are shown with grey-blue surfaces.



Extended Data Fig. 9 | RSB deprotonation pathway in BR. A. The ground state of BR. **B**. The L state of BR. **C**. The M state of BR. H-bonds are shown with black dashed lines. A weak H-bond between T89 and D85 in the L state is shown with black solid line. The distance between RSB and D85 in the L and M states as well as the distance between T89 and D85 in the M state are shown with red arrows. Distance lengths are shown near the lines with bold italic numbers and are in Å. Retinal cofactor and K216 residue are colored teal.



Extended Data Fig. 10 | Bioinformatic analysis of archaeal outward proton-pumping rhodopsins. A. The phylogenetic tree of archaeal outward proton-pumping rhodopsins. BR is shown with red dot. The group of 9 unusual rhodopsins having substitutions at the key functional positions (E194, E204, S193) compared to the other members of the group is contoured and labelled. Also, one of them (E194G, 5 of 9 proteins have this substitution) was shown not to have proton pumping activity. **B**. Amino acid alignment of the representatives of the group of archaeal proton pumps. Only key regions of the protein are shown. Last 4 proteins (marked with a black bracket) belong the group of 9 unusual proteins contoured in panel A. **C**. Fractions of proteins with specific amino acid at specific position according to the BR numeration. The group of 9 unusual rhodopsins is excluded for the calculation of the fractions. Total number of rhodopsins used is 261. **D**. Fractions of proteins with specific amino acid at specific position according to the BR numeration. All 270 proteins were used for calculations. Red arrows indicate essential residues forming the proton wires in course of the BR photocycle.

Supplementary Notes

Supplementary Note 1. The ground state structure of BR.

Compared to the most commonly used model of the BR ground state[37], true-atomic-resolution structures presented here provide additional important features of the cytoplasmic part of the protein essential for proton uptake and reprotonation of RSB (Fig. 1A). First, an additional water molecule is placed in the hydrophobic cavity near W182 (w503, Fig. 1B). It is bound to the T178 side chain and located close to earlier identified water molecule w501 (Fig. 1B). Another water molecule (w502) in this region is placed between RSB and D96 (Fig. 1A). Thus, in contrast to the existing paradigm claiming a large hydrophobic gap between the RSB and its proton donor D96 amino acid[20,46], it appears that even in the ground state, this cytoplasmic region is moderately hydrated.

Second difference is in the PRG consisting of Y83, S193, E194-E204 pair and also involving R82 and surrounding water molecules in the ground state. Namely, for PRG a triple conformation was identified, which is characterized by the presence of very short H-bonds (conformations are indexed with A, B, C; Fig. 1D, 2A; Extended Data Fig. 5C). These multiple conformations were never structurally shown for BR, although a double conformation was observed recently for the ground state of the protein[54] (Extended Data Fig. 5D). In the present model, there are two principally different configurations of the E194-E204 pair; one corresponds to the conformation A, while the second corresponds to the conformations B and C. The conformation A is quite similar to that obtained previously[54] (Extended Data Fig. 5D,E,F); in contrast, the two other conformations are notably different. An additional water molecule not observed previously (w408) interacts directly with E204 in both conformations B and C (Fig. 1D, 2A, 3A). The major differences between conformation B and C are the orientation of S193 residue and positions of water molecules w403 and w404.

Third, our data provide evidence that R82 is already dynamic in the ground state of the protein. This is clearly manifested by the increase of B-factors of the R82 side chain atoms compared to the mean value for nearby residues (21 and 16 Å^2, respectively). The relatively weak and diffusive electron density maps near the NH1, NH2, and CZ atoms of the residue (Fig. 1C) also demonstrate the dynamics of R82. A similar feature was recently observed for BR in ref [54], however, it was not discussed by the authors. Moreover, similar observation was reported recently for another archaeal light-driven proton pump archaerhodopsin-3 (AR3)[67]. These new

findings on the conformational heterogeneity and dynamics of R82 of the ground state of BR are critical for the understanding of proton storage and release mechanisms of BR.

Supplementary Note 2. Evolution of the retinal, RSB, and cytoplasmic part of BR.

Upon absorption of a photon, retinal isomerizes from all-*trans* conformation in the ground state to 13-*cis* conformation in the M state. The isomerization triggers conformational changes of the protein, which enable the vectorial proton transport against the electrochemical gradient. Our data precisely show how the retinal conformation is transformed stepwise through torsional motion and bond stretching along the C_{13}-C_{14} axis from the all-*trans* in the ground state to 13-*cis* conformation in the M state (Fig. 2A). Knowledge of the precise retinal conformation is necessary for understanding the mechanism of the transformation of the energy of the absorbed photon to elastic and electrostatic energy of the chromophore and then to the protein to drive proton transport against the electrochemical gradient. Our crystallographic data provides a unique opportunity to determine *ab initio* the precise conformation of the retinal (Supplementary Table 1).

In the ground state, the structure of the retinal is planar, and the RSB points towards extracellular side, being connected to w402, which is a part of the pentagon of H-bonds formed by two key aspartates (D85 and D212) and three water molecules (w401, w402, and w406) (Fig. 2A,B). In the K state, which appears in picosecond time range, isomerized retinal is slightly bent but the RSB still points towards the extracellular side. This small conformational change leads to break of the H-bond between the RSB and w402 without disturbing the H-bond pentagon (Fig. 2C). Still longer time is required to develop larger amplitude conformational changes.

Progressing to the L state within microseconds, retinal is in not yet fully-relaxed 13-*cis* conformation (Fig. 2A; Supplementary Table 1). However, the RSB already orients itself towards the cytoplasm (Fig. 2A,D). The H-bond pentagon observed in the ground and K states disappears, and two of three water molecules (most likely w402 and w406) disappear from the extracellular environment of the RSB. Presumably, they are relocated upon the K-to-L state transition following the RSB to the cytoplasmic side of the protein (Fig. 2D). Thus, this 'dehydration' and the corresponding breaking of the extensive H-bond network in the extracellular part between the RSB and R82 is accompanied by a synchronous extensive 'hydration' of the cytoplasmic part of the protein between D96 and the RSB. The flip of the L93 side chain during the K-to-L state transition towards the region of W182 is crucial and allows formation of a large cavity in the space between the RSB and D96, filled with five water molecules (Fig. 2D; Fig. 3A; Extended Data Fig. 6). Three of these water molecules mediate

the continuous CHB (proton wire) between the RSB and D96, which stabilizes the protonated form of the RSB in the cytoplasmic part of BR, indicating the presence of the L state (Fig. 2D; Fig. 3A). These data clarify a long-standing controversy between the X-ray and FTIR spectroscopy/molecular dynamics simulations data. The latter, in contrast to previous X-ray crystallography data, claimed the existence of the continuous H-bond network between the RSB and D96 already in the L state[7,47,74]. The RSB is also stabilized by H-bond to T89 in the L state. Two other water molecules in the cytoplasmic cavity stabilize the backbone of the notably distorted helix G in the L state (Fig. 3A; Extended Data Fig. 7). The distortion occurs in the short π-helical region in the central part of the helix G. While in the ground state the α-helical organization of the helix G is disturbed due to the π-bulge at A215[18], in the L-state the backbone of K216 tilts towards the inside of the protein, losing its H-bonding connection to the N atom of the G220 residue (Fig. 3A; Extended Data Fig. 7). This global rearrangement of the helix G, together with the flip of L93, provide enough space to accommodate water molecules mediating H-bond network in the cytoplasmic part of BR in the L state. Importantly, except for the CHB connecting the RSB and D96 in the L state, there is a short branch of the H-bond wire propagating to carbonyl oxygen of K216, also mediated solely by water molecules.

Following this, in the M state, retinal is finally in a relaxed planar 13-*cis* conformation, which is in line with earlier studies[76] (Fig. 2A, Supplementary Table 1). The RSB remains in the cytoplasmic part of the protein stabilized only by T89 (Fig. 2A,E). L93 is reoriented back to its initial ground state position, decreasing the size of the cavity in the cytoplasmic inner part of BR (Fig. 2E; Fig. 3B; Extended Data Fig. 6). The CHB connecting D96 and the RSB in the L state disappears upon L-to-M transition while the second CHB between D96 and carbonyl oxygen of K216 remains but is modified (Fig. 3A,B). Namely, in the M state it involves three water molecules, forming a linear chain of strong short (2.5-2.6 Å) H-bonds between the residues (Fig. 2E; Fig. 3B). Two of these water molecules also stabilize the backbone of the helix G, which remains largely distorted similar to that in the L state (Fig. 3B; Extended Data Fig. 7). We suggest that the switch off of the CHB propagating from D96 to the RSB in the L state prevents the reprotonation of the RSB from its proton donor D96 stabilizing the M state at this stage of the proton translocation process.

Supplementary Note 3. Evolution of the extracellular side of BR.

As it was mentioned above, upon retinal isomerization and consequent RSB reorientation, the H-bond of the RSB to the H-bond pentagon in the extracellular part of BR is already lost in the K state (Fig. 2B,C). However, at this step, the organization of the rest of the extracellular region remains the same as in the ground state (Fig. 2B,C). Namely, the R82 side chain is oriented towards the RSB and is connected to the pentagon (Fig. 2B,C; Fig. 3C). Thus, there are two linear CHBs connecting D85 to R82 in both ground and K states of BR (highlighted violet in Fig. 2B and blue in Fig. 2C). The first one is mediated solely by water molecules w401 and w406. The second one is longer and involves D212 and Y57 side chains and water molecules w402 and w407. D85 is also stabilized by a 2.7 Å H-bond with T89 in the ground and K states.

This picture is changed dramatically upon the K-to-L transition. Water molecules w402, w406, and w407 disappear and, as we have mentioned, are presumably relocated to the cytoplasmic side of BR in the L state (Fig. 2D; Fig. 3A). One water molecule remains bridging D85 and D212. After losing the interactions with w402 upon the K-to-L transition, D212 side chain is stabilized additionally by W86 in the L state. The carboxyl group of D85 is relocated upon the formation of the L state, so the distance between T89 and D85 becomes larger than a typical length of H-bond (Fig. 2D). However, the residues are still interacting over a weak H-bond.

R82, in turn, is reoriented significantly towards the PRG upon the K-to-L state transition (Fig. 2D; Fig. 3D). Interestingly, two water molecules (w403' and w405') appear in the L state in the region between R82, D212, and Y57, stabilizing the residues in the region between the RSB and R82 (Fig. 2D; Fig. 3D). Consequently, two CHBs connecting D85 to R82 in the ground and K states are transformed into a single CHB upon the K-to-L transition (Fig. 2D). This newly formed linear CHB involves D212, Y57, and three water molecules (w401', w403', and w405') (Fig. 2D). Due to the flip of the R82 side chain, the CHB at the extracellular part of BR in the L states terminates at the NE atom of the residue, while in the ground and K states the two CHBs connected D85 directly with the NH1 and NH2 atoms of R82 (Fig. 3C,D).

Importantly, the organization of the region between the RSB and R82 seems to be similar in the L and M states (Fig. 2D,E; Fig. 3D,E). However, the evolution of the retinal configuration and proton translocation from the RSB to D85 upon the L-to-M transition result in several key differences between the structures of the states. They are identified by the high quality of the crystallographic data. First, in the M state the distance between T89 and D85 is enlarged to 4.3 Å. The existence of this gap in the M, but not in the L state indicates that RSB-T89-D85 is the

pathway for proton translocation from the RSB to D85 upon the L-to-M transition, which is broken in the M state to prevent possible proton backflow from D85 to RSB. Second, upon D85 protonation during the L-to-M transition, the length of H-bond between w403' and w405' of the CHB connecting D85 to R82 is changed significantly from 3.0 to 2.4 Å (Fig. 2D,E). Third, R82 is further shifted towards the PRG by 0.3 Å in the M state.

Supplementary Note 4. Conformational changes in the PRG during photocycle

We identified a triple (A, B, and C) conformation of the PRG in the ground state of BR, which is also preserved in the K state. Importantly, several very short H-bonds are present in the A, B, and C conformations of the PRG. First, in the conformation A, the distances between E194 and E204 and between E204 and S193 are 2.4 Å (Fig. 1D, 4A). Second, in the conformations B and C there is a short H-bond between Y83 and E194 (2.4 Å) and between E204 and newly identified water molecule w408 (2.5 Å) (Fig. 4A). Third, in the conformation C there is also a short H-bond between E204 and S193 (2.5 Å) (Fig. 4A). Biological role of the described very short H-bonds in the PRG will be analyzed in the Discussion section of the manuscript.

In contrast, in the L state only a single conformation is observed (Fig. 4B,C). The change in the PRG conformation upon the K-to-L state transition is possibly connected with the flip of the R82 side chain towards the E194-E204 pair. Y83 also shifts by 0.4 Å towards E194. This makes the existence of the conformations B and C, found in the ground and K states, sterically impossible in the L state. The only conformation of the PRG observed in the L state is characterized by very short H-bonds between E194 and E204 (2.3 Å), R82 and E194 (2.4 Å), and E194 and a water molecule w408' (2.5 Å) (Fig. 4B).

In the M state also only a single conformation of the PRG was identified (Fig. 4C). However, it is notably different from that found in the L state. Indeed, when the proton of the PRG is already released, there is no H-bond between E194 and E204 and the E194 is reoriented towards Y83, which stabilizes the glutamate (Fig. 4C). Overall, the structure of the PRG in the M state is similar to some of those obtained earlier[44,51] (Extended Data Fig. 8). However, there is a principal difference. The presented true-atomic-resolution model of the M state allows precise positioning of water molecules and amino acid residues indicating the presence of several very short H-bonds near PRG at the BR extracellular surface (2.4 Å between S193 and w411, also between w411 and carbonyl oxygen of the P77 residue).

Supplementary Note 5. Role of π-bulge in CHBs remodulation and proton transfer.

It has been shown that in the ground state the secondary structure of α-helix G is disrupted by the insertion of the π-bulge at the position of A215[37]. This insertion causes the peptide to kink at this residue which is in direct contact with K216, to which the RSB is attached. The functional role of the bulge is not yet known[37]. Our data show that upon transition to the L state the helix G additionally bends in its central part around K216 towards the inside of the protein. The same part of the helix is slightly further modified in the M state. All this contributes to the switch on of the CHBs propagating from D96 to the RSB and K216 in the L state and its further evolution to a linear water-mediated CHB from D96 to the carbonyl of the K216 amide bond in the M state. We suggest that the latter switch prevents the reprotonation of the RSB from its proton donor D96 stabilizing the M state at this stage of the proton translocation process. Additionally, our data suggest that the rearrangement of the CHBs from the RSB to K216 upon the L-to-M transition reduces pK_a of the RSB triggering its deprotonation. Therefore, the structures reveal in details important functional role of the π-bulge. It facilitates further disturbance of the helix G in the vicinity of K216 and RSB contributing to the rearrangement of H-bonds in the active state of the protein (RSB), which finally leads to its deprotonation.

Supplementary Tables

Supplementary Table 1. Torsion angles of the retinal cofactor of BR in the ground, K, L, and M states.

Angle	Ground state	K state	L state	M state
Cε-Nz=C15-C14	-169.85	161.29	-111.98	-125.22
Nz=C15-C14=C13	173.63	171.77	178.72	-168.80
C15-C14=C13-C12	-154.31	-15.34	17.42	0.02
C14=C13-C12=C11	179.41	-178.89	168.16	-179.98
C13-C12=C11-C10	-170.78	168.91	157.48	167.67
C12=C11-C10=C9	-177.24	-176.69	171.77	173.36
C11-C10=C9-C8	175.34	172.26	-173.50	-179.97
C10=C9-C8=C7	-174.57	180.00	-179.54	179.92
C9-C8=C7-C6	170.39	168.48	165.03	-179.85
C8=C7-C6=C5	178.40	-177.54	-174.89	179.10