nature portfolio

https://doi.org/10.1038/s41594-022-00762-2

Supplementary information

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

In the format provided by the authors and unedited

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