Crystallographic structure of xanthorhodopsin, the light-driven proton pump with a dual chromophore

Hartmut Luecke,a,b,c,1 Brigitte Schobertb, Jason Stagnoa, Eleonora S. Imashevb, Jennifer M. Wangb, Sergei P. Balashovb,c, and Janos K. Lanyi,b,c

Departments of aMolecular Biology and Biochemistry, and bPhysiology and Biophysics, and cCenter for Biomembrane Systems, University of California, Irvine, CA 92697

Edited by Harry B. Gray, California Institute of Technology, Pasadena, CA, and approved September 16, 2008 (received for review July 23, 2008)

Homologous to bacteriorhodopsin and even more to proteorhodopsin, xanthorhodopsin is a light-driven proton pump that, in addition to retinal, contains a noncovalently bound carotenoid with a function of a light-harvesting antenna. We determined the structure of this eubacterial membrane protein–carotenoid complex by X-ray diffraction, to 1.9-Å resolution. Although it contains 7 transmembrane helices like bacteriorhodopsin and archaerhodopsin, the structure of xanthorhodopsin is considerably different from the 2 archaeal proteins. The crystallographic model for this rhodopsin introduces structural motifs for proton transfer during the reaction cycle, particularly for proton release, that are dramatically different from those in other retinal-based transmembrane pumps. Further, it contains a histidine-aspartate complex for regulating the pKa of the primary proton acceptor not present in retinal proteins, bacteriorhodopsin and archaerhodopsin.

Results and Discussion

Xanthorhodopsin was crystallized from bicelles (13), with a type I arrangement of stacked bilayers. The structure was solved to 1.9-Å resolution (Table 1). The P1 unit cell contains 2 molecules of xanthorhodopsin with a head-to-tail arrangement somewhat similar to 2-dimensional crystal forms of bacteriorhodopsin (14) and halorhodopsin (15), as well as 3-dimensional crystals of the D85S bacteriorhodopsin mutant (16), sensory rhodopsin II (17, 18), and Anabaena sensory rhodopsin (19). Considering its function as an ion transporter in the cell membrane, xanthorhodopsin is unlikely to form such dimers in the original Salinibacter cells.

Location and Binding Site of Carotenoid Antenna. The C40 carotenoid lies transverse against the outer surface of helix F at a ≈54° angle to the membrane normal, buried at the protein–lipid boundary (Fig. 2A). Its keto-ring is immobilized by residues at the extracellular ends of helices E and F and by the β-ionone ring of the retinal (Fig. 2B) and rotated 82° from the plane of the methyl side chains of its polyene chain and therefore from the plane of the extended π-system (Fig. 2C). The ring C=O is not hydrogen-bonded. The immobilized and acutely out-of-plane orientation of the keto-ring minimizes participation of its 2 double bonds in the conjugated π-system and explains the well-resolved vibronic bands of the carotenoid, the lack of a red-shift of the absorption bands upon binding, and the strong CD bands in the visible region (5). The relatively rigid polyene is wedged in a slot on the outside of helix F, with one side formed by the Leu-194 and Leu-197 side chains and the other by the Ile-205 side chain, and its base by Gly-201. The carotenoid glucoside is hydrogen-bonded to the C=O and the NH2 of the amidic side chain of Asn-191, as well as NH1 of Arg-184. The increased efficiency of light harvesting by antennae, such as carotenoids, is common in photosynthetic membranes, large multiprotein systems that contain tens or hundreds of chromophores. The much simpler retinal-based archaean proton pumps bacteriorhodopsin and archaerhodopsin lack antennae; light collection and proton translocation are performed by a single protein with a single chromophore. Recently, it was shown that a light-harvesting antenna can function also in a retinal protein, bacteriorhodopsin (1), of the eubacterium (<i>Halobacterium salinarum</i>) (2), contains a single energy-donor carotenoid, salinixanthin (3), and a single acceptor, retinal, in a small (25 kDa) membrane protein. Because energy transfer is from the short-lived S2 carotenoid level (4), there must be a short distance and favorable geometry between the 2 chromophores to account for its high (40–50%) efficiency. Close interaction of the 2 chromophores is indicated by interaction of the carotenoid conformation on the presence of the retinal in the protein (1, 5, 6) and spectral changes of the carotenoid during the photochemical transformations of the retinal (1), but, as for the proteorhodopsin family of proteins, no direct structural information has been available (4, 7). Unexpectedly, the crystallographic structure of xanthorhodopsin we report here reveals not only the location of the antenna but also striking differences from the archaenal retinal proteins, bacteriorhodopsin and archaerhodopsin.

The photocycle of xanthorhodopsin (8) and the functional residues in the ion transfer pathway (1) are similar to those of the numerous other eubacterial proton pumps, the proteorhodopsins (9, 10). Proteins homologous to xanthorhodopsin were found recently in the genome of an abundant coastal ocean methylo troph (11) and earlier in the genomes of <i>Gloebacter violaceus</i>, <i>Pyrocystis lumula</i> (12), and others. The proteins in this clade exhibit significantly less homology to the proteorhodopsins (11). For example, 137 residues (50%) are identical in <i>Gloebacter</i> rhodopsin and xanthorhodopsin, but only 60 residues (22%) in proteorhodopsin and xanthorhodopsin. Although considerable sequence differences separate xanthorhodopsin from the proteorhodopsins (Fig. 1), its structure, the first for a eubacterial proton pump, is likely to be relevant to other eubacterial retinal-based pumps.

Author contributions: S.P.B. and J.K.L. designed research; B.S., E.S.I., J.M.W., S.P.B., and J.K.L. performed research; E.S.I., J.M.W., and S.P.B. contributed new reagents/analytic tools; H.L., J.S., and J.K.L. analyzed data; and H.L., S.P.B., and J.K.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 3DDO).

1To whom correspondence may be addressed. E-mail: hudel@uci.edu, balashov@uci.edu, or jkanyi@uci.edu.

© 2008 by The National Academy of Sciences of the USA
dependence of the carotenoid spectrum on the retinal is explained by the fact that the retinal \( \beta \)-ionone ring is part of the carotenoid binding site (Fig. 2B).

The keto-ring of the carotenoid is in the space occupied by Trp-138 in bacteriorhodopsin, one of the bulky side chains that confines the retinal \( \beta \)-ionone ring in that protein. In xanthorhodopsin, it is replaced by a glycine. This residue replacement might be the best diagnostic in a sequence for the possibility of binding a salinixanthin-like carotenoid. Another difference is Glu-141 of xanthorhodopsin, which is alanine in bacteriorhodopsin but a conserved glutamate in proteorhodopsins and involved in spectral tuning (20), whose carboxyl oxygens are 4 Å from the retinal \( \beta \)-ionone ring methyls. An intriguing question is whether other retinal proteins might also have an antenna. Of the 12 residues in the xanthorhodopsin carotenoid binding site, 7 are conserved in \textit{Gloeobacter} rhodopsin (homologs of Gly-156, Thr-160, Asn-191, Leu-197, Ile-205, Tyr-207, and Met-211). Thus, it is probable that this protein can bind the C\(_{40}\) carotenoid of \textit{Gloeobacter}, echinenone, which contains a keto-ring similar to salinixanthin (21).

**Table 1. Data collection and refinement statistics**

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Beamlne 9.1, SSRL, Menlo Park, CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength, Å</td>
<td>0.979</td>
</tr>
<tr>
<td>Space group</td>
<td>P1</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>( a = 52.7 ) Å, ( b = 59.5 ) Å, ( c = 59.7 ) Å</td>
</tr>
<tr>
<td>( \alpha = 76.4^{\circ} ), ( \beta = 74.9^{\circ} ), ( \gamma = 64.1^{\circ} )</td>
<td></td>
</tr>
<tr>
<td>Resolution range, Å</td>
<td>45.10 – 1.90</td>
</tr>
<tr>
<td>Total reflections</td>
<td>166,560</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>46,289</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.6 (3.5)*</td>
</tr>
<tr>
<td>Completeness, %</td>
<td>94.1 (85.5)*</td>
</tr>
<tr>
<td>Mean ( I/\sigma(I) )</td>
<td>8.4 (1.5)*</td>
</tr>
<tr>
<td>( R_{sym} ), %</td>
<td>5.7 (48.5)*</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
</tr>
<tr>
<td>Resolution range, Å</td>
<td>45.10 – 1.90</td>
</tr>
<tr>
<td>No. of reflections used</td>
<td>46,278</td>
</tr>
<tr>
<td>( R_{free}/R_{work} ), % ( \dagger )</td>
<td>24.7/26.5</td>
</tr>
<tr>
<td>Rmsd bonds, Å</td>
<td>0.012</td>
</tr>
<tr>
<td>Rmsd angles, °</td>
<td>1.35</td>
</tr>
<tr>
<td>No. of atoms/avg. B, Å(^2)</td>
<td>3,925/44.2</td>
</tr>
<tr>
<td>Protein</td>
<td>62/53.2</td>
</tr>
<tr>
<td>Retinal</td>
<td>40/33.4</td>
</tr>
<tr>
<td>Salinixanthin</td>
<td>140/73.0</td>
</tr>
<tr>
<td>Lipids</td>
<td>250/67.4</td>
</tr>
<tr>
<td>Ramachandran plot (favored/allowed/generously allowed), % ( \ddagger )</td>
<td>96/3.3/37.0</td>
</tr>
</tbody>
</table>

\( \dagger \)Values in parentheses are for the highest-resolution shell.

\( \ddagger \)\( R_{free} \) based on a test set size of 4.7%.

\( \ddagger \)PROCHECK

**Fig. 1.** Sequence alignment of green light-absorbing proteorhodopsin (PR), xanthorhodopsin (XR), and bacteriorhodopsin (BR), reevaluated from the one shown in ref. 1 by using information gained from the diffraction structure. Red, conserved residues in all three; purple, conserved residues in xanthorhodopsin and bacteriorhodopsin; yellow, conserved residues in xanthorhodopsin and proteorhodopsins; blue, residues involved with carotenoid binding. Top row of numbers refer to the xanthorhodopsin sequence; bottom row to the bacteriorhodopsin sequence. Underlining indicates residues in transmembrane helices. Proteorhodopsin sequence refers to a species from Monterey Bay, MBP1 (protein accession No. AAG10475).

**Fig. 2.** Location of salinixanthin (orange) and retinal (magenta) in xanthorhodopsin. (A) The extended carotenoid is tightly bound on the transmembrane surface of xanthorhodopsin, traversing nearly the entire bilayer, with an inclination of 54° to the membrane normal. Its keto-ring binds in a pocket between helices E and F, very near the \( \beta \)-ionone ring of the retinal. The angle between the chromophore axes is 46°. The angle between the planes of their \( \pi \)-systems is 68°. Horizontal lines indicate the approximate boundaries of the lipid bilayer. Helices E and F are marked. (B) The binding pocket of the keto ring is formed by Leu-148, Gly-156, Phe-157, Thr-160, Met-208, and Met-211, as well as the retinal \( \beta \)-ionone ring. (C) The keto-ring of the carotene is rotated 82° out of plane of the salinixanthin-conjugated system and is in van der Waals distance of the retinal \( \beta \)-ionone and the phenolic side chain of Tyr-207.
The distance between the centers of the 2 linear chromophores is 11.7 Å. The ring moieties are within 5 Å of one another, and part of the retinal β-ionone ring is in van der Waals distance of the carotenoid keto-ring. Both are in contact with the Tyr-207 ring between them (Fig. 2B and C). This contrasts with the crystal structure of archaerhodopsin (22), a proton pump with a bacterioruberin carotenoid without antenna function (23), where the corresponding center-to-center interchromophore distance is 17 Å, with the closest approach of bacterioruberin to the retinal at 12 Å. In that protein, the carotenoid may have a structural (22) and possibly photoprotective function.

The angle between the axes of the chromophores is 46° (Fig. 2A), somewhat less than the 56 ± 3° estimated from the polarization anisotropy of retinal fluorescence (4). The discrepancy may originate from the off-axis orientation of the transition moment, as in rhodopsin (24). Energy transfer is optimal when the chromophores are parallel. However, where the 2 spectra overlap, the probability of absorbing incident light of all polarization angles is increased when the 2 chromophores are not parallel. This gain would be optimal if the angle between the chromophores were 90°, but that would preclude energy transfer. Thus, the 46° angle in xanthorhodopsin appears to be a compromise between optimal chromophore interaction and optimal collection of light of all polarizations.

In light-harvesting complexes of photosynthetic bacteria, it is the large number of antenna molecules with various orientations (25, 26) that ensures the capture of light of different polarizations.

**Structure of Xanthorhodopsin Exhibits Large Differences from Bacteriorhodopsin.** The xanthorhodopsin structure extends the architecture of the growing family of rhodopsins (27). Remarkably, there are greater differences from the main chain of bacteriorhodopsin than in any of the crystallized microbial rhodopsins, including halorhodopsin, archaerhodopsin, sensory rhodopsin II, and *Anabena* sensory rhodopsin. Helices A and G are longer by 4 and 9 residues, respectively, and their tilt and rotation, particularly of helix A, are considerably different (Fig. 3A). The 28 residues that comprise helix B are 4 residues shifted in the sequence toward the C terminus (i.e., toward the extracellular side). In bacteriorhodopsin, the interhelical B–C antiparallel β-sheet interacts with the D–E loop, whereas in xanthorhodopsin it reorients dramatically to interact with the Arg-8 peptide C=O near the N terminus, where it forms a mini 3-stranded β-sheet. As a result, the tip of the B–C loop is displaced, by 30 Å, toward the periphery of the protein (Fig. 3). Unexpected in a heptahelical membrane protein, this produces a large cleft that extends far into the interior and brings functional residues, buried in other rhodopsins, near the aqueous interface (Fig. 3B).

In bacteriorhodopsin, Wat-402 receives a hydrogen bond from the protonated retinal Schiff base and donates hydrogen bonds to the 2 anionic residues Asp-85 and Asp-212 (28, 29), and this arrangement is conserved in xanthorhodopsin (Fig. 4). However, the carboxylate of the homolog of Asp-85, Asp-96 in this protein, is severely rotated, and the hydrogen-bonded aqueous network of water molecules in the extracellular region that facilitate proton release in the photocycle (30) is replaced by hydrogen-bonded residues that are likely to be more resistant to rearrangement than an aqueous network. In bacteriorhodopsin, a pair of glutamate residues, Glu-194 and Glu-204, stabilizes a hydrogen-bonded aqueous network (31), which is the source of the proton released to the extracellular surface after the retinal Schiff base is deprotonated. Sequence alignment (Fig. 1) shows that the cubacterial pumps, proteorhodopsin and xanthorhodopsin, contain only 1 of these acidic groups. Further, in xanthorhodopsin, at least, the single glutamate is far removed from Arg-93 (>18 Å vs. 7.3 Å in bacteriorhodopsin). In bacteriorhodopsin, release of the proton is triggered by the movement of the positively charged arginine (Arg-82 in this protein) side chain toward the glutamate pair, upon protonation of Asp-85 (30). This is less likely to occur in the xanthorhodopsin photocycle, because NH1 and NH2 of Arg-93 are both hydrogen-bonded to the peptide carbonyl of Gln-229 instead of water molecules.

**Fig. 3.** Displacements of the B–C and F–G interhelical segments expose a deep cavity that extends from the extracellular side halfway toward Schiff base. (A) Comparison of xanthorhodopsin and bacteriorhodopsin. The antiparallel β-sheet of the B–C segment of xanthorhodopsin (gray) packs against the N terminus (Bottom Right), whereas in bacteriorhodopsin (magenta) this segment packs against the F–G loop (Bottom Left). Helices A–E are viewed from the front; F and G are in the back, as marked. (B) The resulting hydrophilic cavity in xanthorhodopsin extends from the extracellular surface to Arg-93 and other buried functional groups. In bacteriorhodopsin, this region is occupied by the protein and includes the proton release group composed of Glu-194, Glu-204, and 3 ordered water molecules, absent in xanthorhodopsin. Buried residues with functional roles in transport are shown to illustrate their proximity to the aqueous interface.

**Fig. 4.** Structure of the retinal, the Schiff base counterion, and the extracellular region. The counterion to the Schiff base is an aspartate–histidine complex. The network of water molecules that leads to the extracellular surface in bacteriorhodopsin is missing, and Arg-93 interacts primarily with protein side chains.
bacterial pumps. It is likely to account not only for the observed orhodopsins (see Fig. 1 and ref. 37), making it likely that the proton to yield a cationic complex.

acceptor, because the carboxylate may accommodate another ionic, rather than the neutral complex (Fig. 5), is the proton counterion. If analogy with bacteriorhodopsin holds, the an-

protonation states of the aspartate–histidine counterion complex. The pH of the crystallization (5.6), is below the observed (8) spectral transition between the protonated and deprotonated forms of the Schiff base counterion (presumably neutral/zwitterionic and anionic). Thus, it seems likely that the crystallographic structure contains the neutral/zwitterionic counterion.

An Asp–His Hydrogen-Bonded Pair Is Part of Counterion to Retinal Schiff Base. One of the distinguishing features of eubacterial proton pumps is that the pH of the primary proton acceptor is not as low as 2.5 in bacteriorhodopsin, but near 7 (32, 33). The origin of the high pH, which makes these proteins functional as pumps only at alkaline pH (34), has been an unsolved problem. In xanthorhodopsin, ND1 of His-62 is hydrogen-bonded to OD1 of Asp-96; at 2.42/2.55 Å (in the 2 molecules of the asymmetric unit), this is a very short hydrogen bond, such as found at active sites of numerous proteins (35). Thus, its proton may be shared by the imidazole ring and the carboxylate in a single-well, strong hydrogen bond, and the complex, with an expected pKa higher than the aspartate alone (36), must be regarded as the Schiff base counterion. If analogy with bacteriorhodopsin holds, the anionic, rather than the neutral complex (Fig. 5), is the proton acceptor of the Schiff base in the photocycle. However, we cannot exclude the possibility that the neutral form is the proton acceptor, because the carboxylate may accommodate another proton to yield a cationic complex.

A histidine at this position is highly conserved in the prote-

Fig. 6. Cytoplasmic region, with the proton donor Glu-107 and its link via Wat-502 to the retinal region. As in the other microbial rhodopsins, Wat-501 is hydrogen-bonded tightly between the tryptophan just above the retinal (Trp-200) and the main-chain carbonyl of the residue in the π-bulge of helix G (Ala-239). Helices B, C, F, and G are marked.

orhodopsins (see Fig. 1 and ref. 37), making it likely that the aspartate–histidine complex is a general characteristic of eubacterial pumps. It is likely to account not only for the observed high pH, of the counterion in these retinal proteins, but also for the smaller red shift of the chromophore maximum upon the protonation of the counterion in xanthorhodopsin than in bacteriorhodopsin (3–5 nm vs. 40 nm) (8) because sharing of the proton with the histidine would leave a partial charge on the aspartate ring.

Once protonated in the photocycle, the His-62–Asp-96 com-

proton release complex (38, 39) and as in proteorhodopsin (32), there is no release of a proton in the xanthorhodopsin photocycle at the time the Schiff base becomes deprotonated. Instead, the sequence of proton release to the extracellular side and uptake from the cytoplasmic side is reversed; proton uptake in the cycle occurs first, evidently by Glu-107 after it has reprotonated the retinal Schiff base, and the release to the extracellular surface is delayed until the final photocycle step. The rationale for keeping counterion protonated until the end of the photocycle might be the same as in bacteriorhodopsin for which it was demonstrated (40) that a neutralized counterion facilitates thermal retinal isomerization, an obligatory step in the reaction cycle.

Hydrogen-Bonded Aqueous Network in the Cytoplasmic Domain Needs Less Rearrangement for Proton Transfer. In the cytoplasmic region of bacteriorhodopsin, the proton donor is in an anhydrous environment that constitutes the hydrophobic barrier in the cytoplasmic half of the protein. This, and the fact that OD1 of Thr-46 is an acceptor of its proton (28), raises its pKa. The aspartic acid becomes a proton donor to the Schiff base during the photocycle only after the Glu–Thr hydrogen bond is broken and this region is hydrated so as to create a hydrogen-bonded chain of 4 water molecules to connect the proton donor to its acceptor (41). In xanthorhodopsin, as in the proteorhodopsins, these residues are replaced by a glutamic acid and a serine (Fig. 1), and they are not hydrogen-bonded to one another. The carboxyl is hydrogen-bonded to Wat-502 that connects to the peptide carbonyl of Lys-240 (Fig. 6). It appears, therefore, that in xanthorhodopsin part of the cytoplasmic hydrogen-bonded chain of water molecules between the retinal and the proton donor is in position already before photoisomerization of the retinal.

As in the other structurally characterized microbial rhodopsins, the other cytoplasmic water, Wat-501, is hydrogen-bonded tightly between the tryptophan located immediately on the cytoplasmic side of the retinal (Trp-200 in xanthorhodopsin) and the main-chain carbonyl of the residue at the apex of the π-bulge of helix G (Ala-239), a connection that is broken during the photocycle as the movement of the retinal C20 displaces the side chain of Trp-200 by >1 Å (30). Wat-501 possesses one of the lowest B factors of any atom in the model, and it is noteworthy that its presence appears to induce a deviation from planarity in the nearby tyrosine side chain.

Materials and Methods

A membrane fraction enriched in xanthorhodopsin was prepared by washing S. ruber cell membranes 4 times with distilled water, followed by washing 3
times with 0.01% dodecyl maltoside in 100 mM NaCl and 5 mM Na,N,N,N-tetrakis(2-hydroxyethyl)glycine (pH 8). The resulting membrane fraction was resuspended in 30 mM phosphate (pH 5.6) and 1 mM sodium azide and concentrated to contain 5 mg/ml xanthorhodopsin (as estimated from absorption at 560 nm). The protein was solubilized in a bicelle-type medium (13) by adding 1 M of solubilized xanthorhodopsin, 3 M of 3 M sodium phosphate (pH 5.6), and 2 M of 2.5 mM sodium azide. After 5 min-equilibrations with 5%, 10%, and then 15% ethylene glycol, the crystals were frozen rapidly in liquid nitrogen. Data were collected at 100 K on beamline 9.1 at the Stanford Synchrotron Research Laboratory (SSRL) as 360 frames with 1° rotations. Data reduction statistics are listed in Table 1.

The structure was solved by iterative molecular replacement, using a composite model that consisted of helices A and B of Anabaena sensory rhodopsin [Protein Data Bank (PDB) ID code 1XIO, residues 81–231] and helices C–G of bacteriorhodopsin (PDB ID code 1C3W, residues 301–505) with the program Phaser (42). The first rotation function exhibited high signal-to-noise ratio with a top Z-score of 4.77. The correct solution was peak 5 with a Z-score of 4.73. After 10 cycles of restrained refinement with the program REFMAC (43), 1 molecule of the resulting model was used for a second round of molecular replacement, yielding much-improved signal-to-noise ratio with Z-scores of 7.32 and 7.10 for the two rotation functions, respectively. Maps were improved by 2-fold averaging with the program DM (44). Electron density for the salinixanthin carotenoid was readily visible at this stage. The model was improved by iterative cycles of model building and refinement with the program COOT (45) and refinement with the program CNS (46), initially with and later without NCS restraints. The rmsd values between the final model (molecule A/B) and the search model, bacteriorhodopsin and Anabaena sensory rhodopsin, are 1.91/1.88 Å and 2.00/1.98 Å, respectively. Sequence identities with bacteriorhodopsin and Anabaena sensory rhodopsin are 24.6% and 21.5%, respectively. In addition to the carotenoid, numerous lipids and lipid fragments have been included in the model. All residues fall within the allowed regions of the Ramachandran plot. Refinement statistics are listed in Table 1. The coordinate file, 3DLI, has been deposited at the Protein Data Bank.

ACKNOWLEDGMENTS. We thank T. Poulos and H. Li (University of California, Irvine) for sharing transportation of crystals and beam time and for help with remote data collection. This work was supported by National Institutes of Health Grant GM29498 and Department of Energy Grant DEFG03-86ER13525 (to J.K.L.), U.S. Army Research Office Grant W911NF-06-1-0020 (to S.P.B.), and National Institutes of Health Grant GM56445 (to H.L.).