

A Novel Membrane Protein in the *Rhodobacter sphaeroides* LH1-RC Photocomplex

K. Tani^{1*}, K. V. P. Nagashima², R. Kanno³, S. Kawamura⁴, R. Kikuchi⁴, M. Hall³, L.-J. Yu⁵, Y. Kimura⁶, M. T. Madigan⁷, A. Mizoguchi¹, B. M. Humbel³ & Z.-Y. Wang-Otomo^{4*}

¹ Graduate School of Medicine, Mie University, Tsu 514-8507, Japan

² Research Institute for Integrated Science, Kanagawa University, 2946 Tsuchiya, Hiratsuka, Kanagawa 259-1293, Japan

³ Imaging Section, Research Support Division, Okinawa Institute of Science and Technology Graduate University (OIST), 1919-1, Tancha, Onna-son, Kunigami-gun, Okinawa 904-0495, Japan

⁴ Faculty of Science, Ibaraki University, Mito 310-8512, Japan

⁵ Photosynthesis Research Center, Key Laboratory of Photobiology, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China.

⁶ Department of Agrobioscience, Graduate School of Agriculture, Kobe University, Nada, Kobe 657-8501, Japan

⁷ School of Biological Sciences, Department of Microbiology, Southern Illinois University, Carbondale, IL 62901 USA

*Corresponding authors, Email: ktani@doc.medic.mie-u.ac.jp; wang@ml.ibaraki.ac.jp

Abstract: We present a cryo-EM structure of the monomeric light-harvesting-reaction center (LH1-RC) core complex from photosynthetic purple bacterium *Rhodobacter (Rba.) sphaeroides* at 2.9 Å resolution. The LH1 complex forms a C-shaped structure composed of 14 $\alpha\beta$ -polypeptides around the RC with a large ring opening. From the cryo-EM density map, a previously unrecognized integral membrane protein, referred to as protein-U, was identified. Protein-U has a U-shaped conformation near the LH1-ring opening and was annotated as a hypothetical protein in the *Rba. sphaeroides* genome. Deletion of protein-U resulted in a mutant strain that expressed a much-reduced amount of the dimeric LH1-RC, indicating an important role for protein-U in dimerization of the LH1-RC complex. PufX was located opposite protein-U on the LH1-ring opening, and both its position and conformation differed from that of previous reports of dimeric LH1-RC structures obtained at low-resolution. Twenty-six molecules of the carotenoid spheroidene arranged in two distinct configurations were resolved in the *Rba. sphaeroides* LH1 and were positioned within the complex to block its pores. Our findings offer a new view of the core photocomplex of *Rba. sphaeroides* and the connections between structure and function in bacterial photocomplexes in general.

Rba. sphaeroides has been the most thoroughly investigated purple phototroph and has been widely used as a model for fundamental biochemical genetic studies of photochemistry, metabolism and regulation. A crystal structure of the reaction center (RC) complex from *Rba. sphaeroides* was the second such to appear^{1,2} following shortly after that of the RC from *Blastochloris (Blc.) viridis*³. The two structures displayed high similarities in the arrangements of most proteins and cofactors, suggesting that the core components of the RC were structurally conserved among all purple bacteria⁴.

The core light-harvesting (LH1) system of *Rba. sphaeroides* is unusual. The LH1 complex composed of α - and β -polypeptides forms an S-shaped structure in the native membranes^{5,6} and 2D crystals⁷⁻⁹. The S-shaped LH1 incorporates two RCs resulting in a so-called dimeric LH1-RC complex, a structure that was also observed in *Rba. blasticus*¹⁰ and *Rhodobaca bogoriensis*¹¹, and was suggested to exist in a few other purple bacteria¹². A key component thought to be responsible for the dimerization is PufX, a protein present within the LH1-RC core complex^{13,14}. PufX is present in all species of *Rhodobacter*¹⁵ and has been shown to be essential for phototrophic growth in both *Rba. sphaeroides* and *Rba. capsulatus*¹⁶⁻¹⁸. Previous work has probed the structural and functional role of PufX and implicated this key protein in regulation of membrane morphology, core photocomplex organization, and cyclic electron transfer¹⁹.

In addition to the dimeric core complex, a monomeric LH1-RC-PufX complex also exists in native membranes of *Rba. sphaeroides* and *Rba. blasticus* and can be separately purified from the dimeric form^{7,10,12}. The monomeric core complex is identical to one-half of a dimer with a gap in the LH1 ring, displaying a C-shaped arrangement¹⁰. It was demonstrated that monomers can reform dimers upon reconstitution⁷, suggesting that the monomer represents a native state in equilibrium with the dimeric complexes. Indeed, only the monomeric form has been observed for the core complexes from *Rba. veldkampii*^{20,21} and *Rba. capsulatus*¹². In addition to its structure, *Rba. sphaeroides* LH1 is also unique in its carotenoid composition. Regardless of dimeric or monomeric form, the *Rba. sphaeroides* LH1 complex contains approximately two carotenoids per $\alpha\beta$ -polypeptides²²⁻²⁴, the highest ratio among known LH1 complexes of purple phototrophs. However, despite the likely importance of carotenoids in LH1 function, *Rba. sphaeroides* LH1 carotenoids have not been resolved from low-resolution structural analyses (an 8-Å structure is the best thus far attained)²³; thus, the precise position and function of carotenoids in the *Rba. sphaeroides* LH1 remains unknown.

Here we present a cryo-EM structure of the monomeric core complex from *Rba. sphaeroides* f. sp. *denitrificans* (strain IL106)²⁵ at 2.9 Å resolution (Fig. 1, Supplementary Table 1 and Supplementary Figs. 1–4). The LH1 complex is composed of 14 pairs of helical $\alpha\beta$ -polypeptides, 28 BChls *a* and 26 spheroidenes surrounding the RC with a large opening in the slightly elliptical C-shaped ring structure. Although structures of the LH1 $\alpha\beta$ -polypeptides and all proteins in the RC are consistent with previous work, it is notable that PufX was found at a different position and with a different conformation from that in the dimeric structure (PDB: 4JC9, 4JCB) determined at 8 Å for the *Rba. sphaeroides* LH1-RC (strain DBCΩG, Supplementary Fig. 5)²³. In addition, a large fragment of electron potential densities near the LH1-ring opening was observed in the refined map but could not be modeled by any known proteins or cofactors. Because our density map at 2.9 Å resolution was of high quality, we were able to trace not only the mainchain of this protein but also the sidechains for many residues. The derived amino acid sequence allowed us to search the *Rba. sphaeroides* IL106 genome (GenBank assembly accession: GCA_003363065.1), and in doing so, a hypothetical protein (DWF04_22265) containing 53 amino acids emerged whose sequence turned out to perfectly fit the density map (Fig. 2b). Hereafter, we refer to this previously unrecognized protein as “protein-U”, the seventh distinct protein in the *Rba. sphaeroides* LH1-RC complex (Fig. 1c, 1d; Fig. 2a, 2b).

The *Rba. sphaeroides* LH1 $\alpha\beta$ -polypeptides form a roughly elliptical C-shaped arrangement (Fig. 1b). The LH1 α -polypeptides are formylated at their N-terminus as confirmed by TOF/MS (Supplementary Fig. 6) and the cryo-EM density map (Supplementary Fig. 4). These formyl groups were modeled in the LH1 structure and play a role in stabilizing the N-terminal regions of the α -polypeptides through hydrogen bonding. PufX and protein-U are located opposite each other on the two sides of the LH1-ring opening (Fig. 1, Fig. 2a). As

a result, the positions of several LH1 $\alpha\beta$ -subunits near protein-U gradually deviated from the closed LH1 ellipse (Fig. 1d) that has been observed in the monomeric type LH1-RCs from *Thermochromatium (Tch.) tepidum*²⁶ and *Rhodopseudomonas (Rps.) palustris*²⁷.

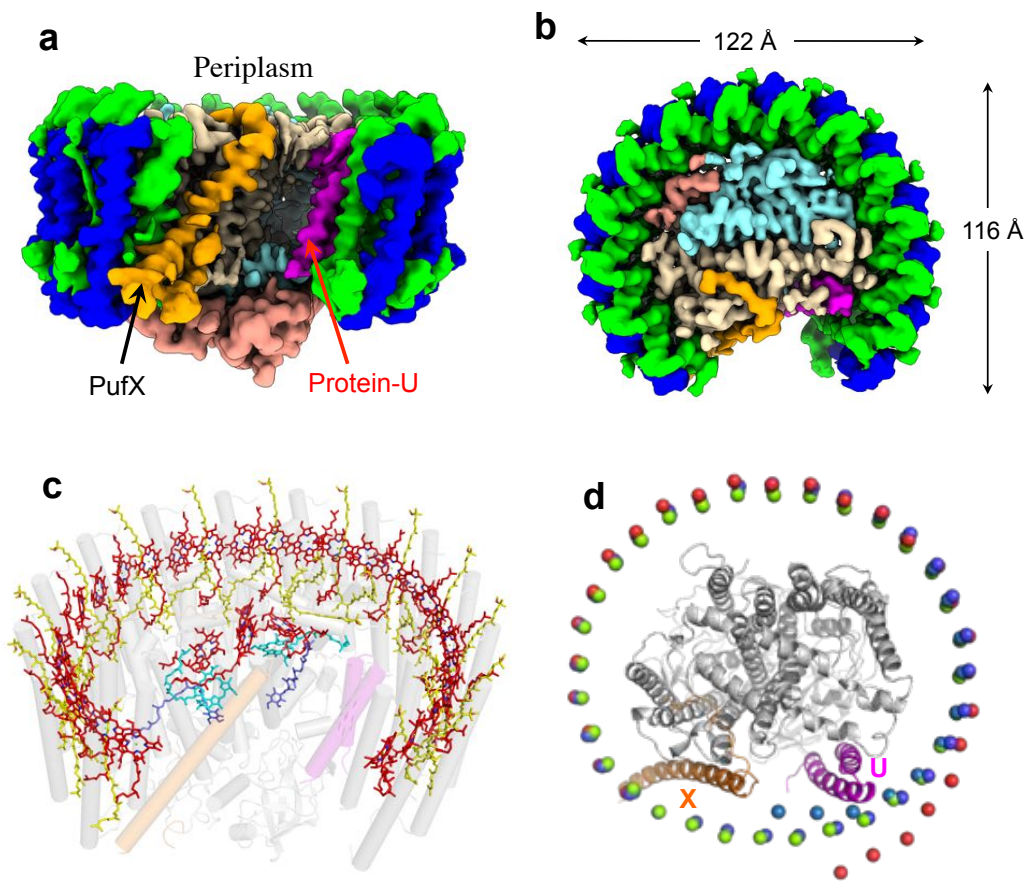


Fig. 1 Structure overview of the monomeric LH1-RC complex from *Rba. sphaeroides* IL106. (a) Side view of surface representations for the LH1-RC proteins parallel to the membrane plane. (b) Top view of surface representations for the LH1-RC proteins from the periplasmic side of the membrane. Color scheme: LH1- α , green; LH1- β , blue; RC-L, wheat; RC-M, cyan; RC-H, salmon; PufX, orange; protein-U, magenta. (c) Tilted view of the cofactor arrangement. Color scheme: BChl *a*, red sticks; spheroidenes, yellow sticks; BPhe *a*, cyan sticks; UQ-10, blue sticks; PufX, transparent orange cylinder; protein-U, transparent magenta cylinder, and transparent gray cylinders for all other proteins. (d) Comparisons of the central Mg atom positions in the LH1 BChl *a* molecules from *Rba. sphaeroides* (red spheres) with those from *Tch. tepidum* (green spheres, PDB: 5Y5S), protein-W-containing *Rps. palustris* (cyan spheres, PDB: 6Z5S), and protein-W-deficient *Rps. palustris* (blue spheres, PDB: 6Z5R). The structures were superimposed by C α carbons of the RC (L, M, H) proteins. The ribbon models represent the RC-L, -M and -H subunits (gray), PufX (orange), and protein-U (magenta) of the *Rba. sphaeroides* IL106.

The Unreported Protein-U. *Rba. sphaeroides* protein-U is located in the interior of the LH1-RC complex between the RC-L subunit and two LH1 α -polypeptides and exhibits a helix-turn-helix conformation with both N- and C-termini on the cytoplasmic side (Fig. 2a, 2b). The two transmembrane (TM) helical regions contain highly hydrophobic residues rich

in Gly and Ala, and are connected by a short loop of five residues. The characteristic U-shaped feature was clearly observed in the cryo-EM map and is consistent with that predicted by the membrane protein topology program TMHMM²⁸ (Fig. 2d). Protein-U interacts with surrounding proteins mainly through its terminal domains and loop region. The mainchain oxygen of protein-U Val4 and sidechain carboxyl group of Glu6 form hydrogen bonds with the sidechain of Arg15 in an α -polypeptide, while a segment (Thr49–Pro50–Asn51) in the protein-U C-terminal region form hydrogen bonds with two α -polypeptides (Fig. 2c). In the loop region of protein-U, Trp32 and Phe33 form close contacts with Trp265 and Trp266 in the RC-L subunit where a putative UQ-10 molecule is present nearby. The unique position and interaction pattern of protein-U indicate that one of its functions may be as a “spacer” pushing several nearby LH1 polypeptides outward from the RC as shown in Fig. 1d and implying a structural role in producing steric hindrances to prevent the formation of a closed LH1 ring.

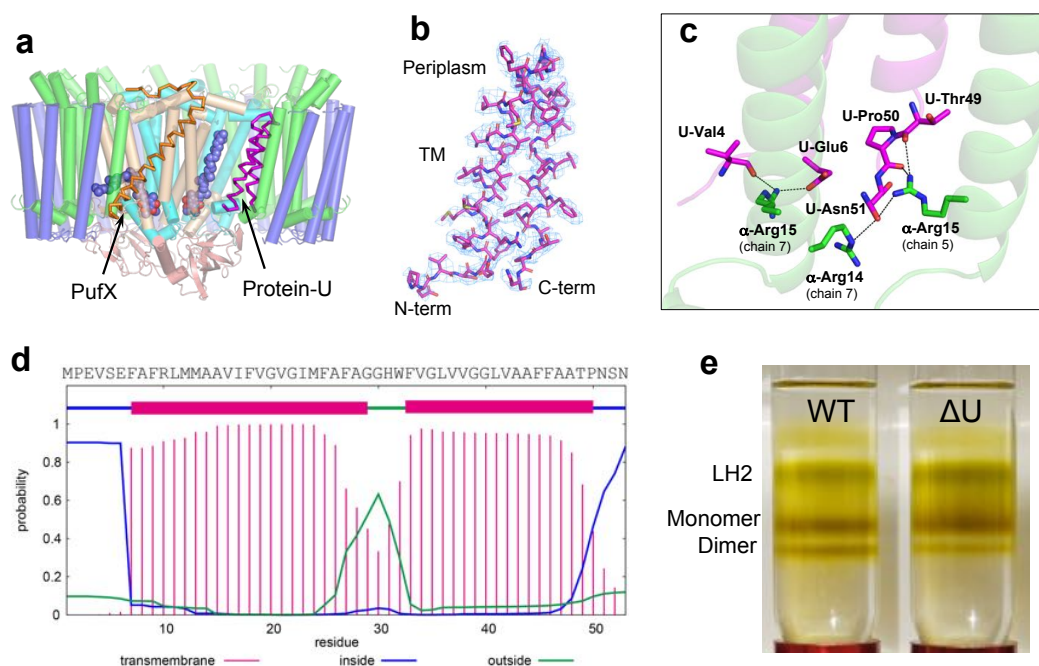


Fig. 2 Structure and property of *Rba. sphaeroides* protein-U. (a) Relative position of protein-U (magenta ribbon) in the LH1-RC. PufX is shown by the orange ribbon and the UQ-10 at Q_A and Q_B sites are shown by blue spheres. All other proteins are shown by cylinders with the same color scheme as in Fig. 1(a). (b) Structure of the protein-U and its density map at a contour level of 3.0 σ . (c) Close contacts (<3.5 Å, dashed lines) between the protein-U terminal regions (magenta) and LH1 α -polypeptides (green). (d) Amino acid sequence of protein-U and its two consecutive transmembrane regions predicted by the membrane protein topology program TMHMM. (e) Sucrose density gradient (10–40% w/v) centrifugations of the solubilized pigment-protein complexes from wild-type (WT) and protein-U-deleted (Δ U) *Rba. sphaeroides* IL106 membranes.

To further investigate potential functions of protein-U, we constructed a mutant strain (IL106- Δ U) deficient in protein-U (Supplementary Fig. 7). This deletion mutant was capable of phototrophic growth at a growth rate similar to that of wild-type cells. However, levels of dimeric LH1-RC significantly decreased in the mutant compared with the wild type. Sucrose density gradient analysis revealed that the ratio of dimeric to monomeric LH1-RC was cut by

nearly two-thirds (0.44 to 0.15) in the mutant (Fig. 2e, Supplementary Fig. 8a), indicating that in addition to a structural function as mentioned above, protein-U likely plays a functional role in forming dimers of the LH1-RC complex. No apparent effects of the absence of protein-U on absorption spectra of either dimeric or monomeric LH1-RCs were detected (Supplementary Fig. 8b).

The PufX Polypeptide. The full-length PufX of *Rba. sphaeroides* IL106 is encoded by 82 amino acids in which the N-terminal Met and C-terminal 13 residues are post-translationally removed in the expressed polypeptide as confirmed by TOF/MS (Fig. 3d, Supplementary Fig. 6). The PufX sequence of strain IL106 differs slightly from that of the type strain 2.4.1 at Asp58 and Pro61, where both are replaced with Gln in the *Rba. sphaeroides* strain 2.4.1. In our cryo-EM structure, the N-terminal 15 residues of the expressed PufX were invisible presumably due to disordered conformations. PufX is located on the opposite side of protein-U in the LH1-ring opening (Fig. 2a) at a position close to that of protein-W found in the *Rps. palustris* LH1-RC²⁷ although the two proteins share no sequence similarity. PufX has a single, unusually long transmembrane domain with its N-terminus on the cytoplasmic side and C-terminus on the periplasmic surface (Fig. 3a). Differing from other proteins in the LH1-RC, PufX displays a remarkably tilted conformation for its transmembrane helical domain with an angle of approximately 52° inclined toward the presumed membrane plane (Fig. 3a, Supplementary Fig. 10); this results in a much longer membrane-spanning region (38 residues) in good agreement with an experimental estimation for the core segment in PufX²⁹.

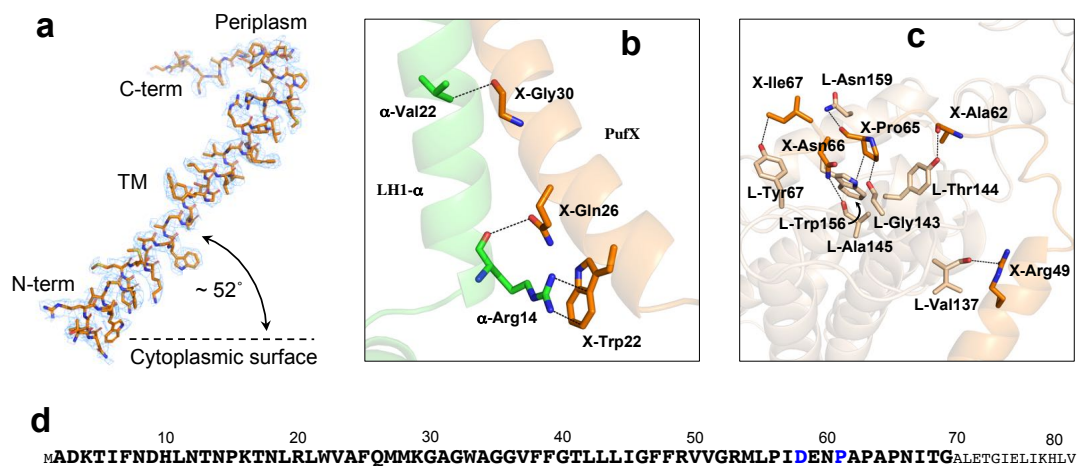


Fig. 3 Structure and interactions of *Rba. sphaeroides* PufX. (a) Structure and density map (3.0 σ) of PufX with orientation toward the presumed membrane plane. (b) Close contacts (<3.5 Å, dashed lines) between the PufX N-terminal region (transparent orange) and an LH1 α -polypeptide (chain ID: 7, transparent green). (c) Close contacts (<3.5 Å, dashed lines) between the PufX C-terminal domain (transparent orange) and the RC L-subunit (transparent wheat). (d) Primary sequence of the full-length PufX in *Rba. sphaeroides* IL106. Sequence for the expressed PufX is shown by larger bold fonts. Membrane-spanning residues are underlined. The Glu58 and Pro61 (blue fonts) are both replaced by Gln in the type strain 2.4.1.

No strong interactions (<4.0 Å) were observed between PufX and pigment molecules (BChl *a* and carotenoids). However, the N-terminal region of PufX interacts with an LH1 α -polypeptide, while the C-terminal domain has closer contacts with the RC L-subunit on the

periplasmic surface. The Trp22 in PufX likely has cation– π interactions with the sidechain of Arg14 in an α -polypeptide (Fig. 3b). The Arg49 and Ala62 in PufX form hydrogen bonds with Val137 and Thr144 in the RC L-subunit, respectively (Fig. 3c). The extensive interactions between the PufX C-terminal end (Ala62–Ile67) and a segment of the RC L-subunit (Gly143–Trp156) explain the experimental observation that the C-terminus of PufX plays an important role in dimerization and assembly of the LH1-RC in *Rba. sphaeroides*³⁰.

The Carotenoids in the *Rba. sphaeroides* LH1 Complex. A total of 26 all-*trans*-spheroidenes in the LH1 complex of *Rba. sphaeroides* IL106 were clearly resolved in our cryo-EM map (Fig. 1c, Fig. 4a, Supplementary Fig. 4). This corresponds to a ratio of approximately two carotenoids per $\alpha\beta$ -polypeptides in contrast to the one carotenoid per $\alpha\beta$ -pair commonly found in other LH1 complexes^{26, 27, 31, 32}. All methoxy groups of the *Rba. sphaeroides* carotenoids point toward the periplasmic side, and the carotenoids can be grouped into two sets based on their orientations and positions in the LH1. Group-A spheroidenes are embedded deeply in the transmembrane region between the α - and β -polypeptides at similar positions and with a similar conformation to those in the LH1 complexes of other purple bacteria (Fig. 4b)^{26, 27, 31-33}. The methoxy groups of these carotenoids are positioned close to the center of bacteriochlorins (Fig. 4a) and there are no apparent associations between the carotenoids and surrounding polypeptides and BChl *a* molecules.

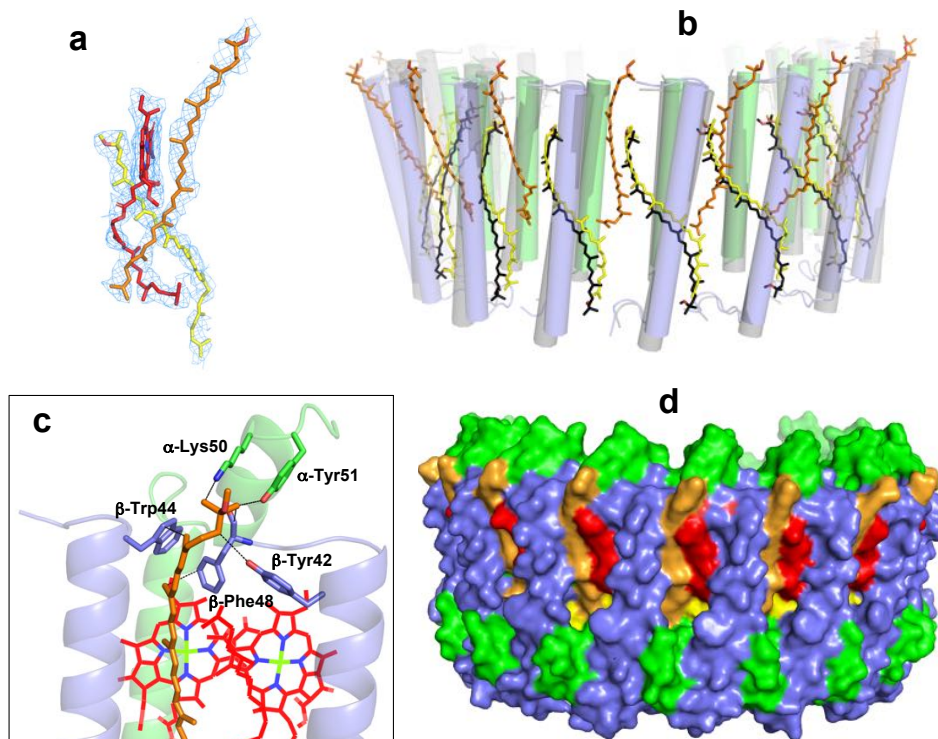


Fig. 4 Carotenoids in the *Rba. sphaeroides* LH1. (a) Typical structures and density maps (3σ) of two spheroidenes (yellow and orange) relative to the position of a BChl *a* (red) in the LH1 complex with different configurations. (b) Two groups of the spheroidenes (yellow and orange sticks) in the *Rba. sphaeroides* LH1 in comparison with the carotenoids in the *Tch. tepidum* LH1 (black sticks, PDB:5Y5S). The structures were superimposed by C α carbons of the LH1 $\alpha\beta$ -polypeptides between *Rba. sphaeroides* (transparent green and slate-blue cylinders) and *Tch. tepidum* (transparent gray cylinders). (c) Close contacts (<4.0 Å, dashed

lines) between a Group-B spheroidene and its surrounding $\alpha\beta$ -polypeptides. **(d)** Side view of the surface representation shows a sealed fence for the *Rba. sphaeroides* LH1 complex. Color scheme: LH1- α , green; LH1- β , slate-blue; BChl *a*, red; spheroidenes, yellow (Group-A) and orange (Group-B).

In contrast to Group-A spheroidenes, Group-B spheroidenes adopt a distinct configuration within the *Rba. sphaeroides* LH1 with a large shift toward the periplasmic side (Fig. 4b). Each of these carotenoids is located in between two adjacent $\alpha\beta$ -pairs with their methoxy group protruded on the periplasmic surface. The methoxy ends of these carotenoids are in close proximity to the C-terminal ends of $\alpha\beta$ -polypeptides (Fig. 4c) and the central portions form hydrophobic interactions with BChl *a* molecules. The additional Group-B spheroidenes, combined with Group-A spheroidenes, fill the space between adjacent $\alpha\beta$ -subunits (Fig. 4d), and this leads to blockage of the pores (channels) observed in BChl *a*-containing LH1 complexes containing only Group-A carotenoids^{26, 27, 32, 34}. The combination of both Group A and B carotenoids results in a tightly sealed, impenetrable LH1 “fence” that requires alternative strategies for facilitating quinone transport between the *Rba. sphaeroides* RC and quinone pools during photosynthetic electron flow. These alternative strategies include the C-shaped monomeric LH1 whose structure was revealed herein or an S-shaped dimeric LH1 with incomplete rings as previously reported²³.

Discussion

The gene encoding the 53 amino acids in protein-U is flanked in the *Rba. sphaeroides* IL106 genome by genes encoding a transglutaminase family protein and metallo-hydrolase and is the only gene in this region transcribed in the opposite direction (Supplementary Fig. 7). This suggests that transcription of the gene encoding protein-U is controlled independently from that of adjacent genes. The expressed protein-U locates in the interior space between LH1 and RC near the edge of the LH1-ring opening. The corresponding position is occupied by a cardiolipin molecule in the protein W-containing *Rps. palustris* LH1-RC that also has an opening in the LH1 ring²⁷. Protein-U has a tandem arrangement of two membrane-spanning segments connected by a short loop, forming a U-shaped conformation. High frequencies of Gly and Ala residues in the transmembrane region indicate a tendency for these α helices to pack tightly at particular sites such as Gly22, Gly35 and Gly39 (Fig. 2d) to form a canonical left-handed supercoil as is often seen in α -bundle membrane proteins³⁵⁻³⁷. Such helical pairs with antiparallel arrangements are known to form large interfacial areas, which in the case of protein-U may contribute to LH1-RC stabilization and/or facilitate dimerization of the complex.

A genomic database search of the genus *Rhodobacter* revealed that six of fifteen species contain protein-U (or protein-U-like) genes, and all *Rba. sphaeroides* strains with sequenced genomes encode this protein (Supplementary Fig. 9a). Proteins-U can be classified into three types based on their sequences (Supplementary Fig. 9b): Type-1 (WP_002721225) for *Rba. johrii*, *Rba. megalophilus* and all but one *Rba. sphaeroides* strains; Type-2 (WP_176504535) for *Rba. ovatus*; and Type-3 (WP_085996593) for *Rba. azotoformans*, *Rba. sediminicola* and *Rba. sphaeroides* ATCC 17025. Type-1 and Type-3 proteins-U have the same number of residues and share high sequence identity, whereas Type-2 has a long N-terminal domain but also shares high sequence identity with Type-1 and Type-3 proteins-U in other portions (Supplementary Fig. 9b). All *Rhodobacter* species containing protein-U also have *pufX* (or *pufX*-like) genes but the converse is not true. Considering that a *Rba. sphaeroides* protein-U-deletion strain in our work and wild-type *Rba.*

blasticus that naturally lacks protein-U¹⁰ are both capable of photosynthetic growth and producing at least some dimeric LH1-RC, protein-U is clearly not indispensable for photosynthesis. Therefore, we conclude that a major function of protein-U is to increase dimerization of the LH1-RC complex from that would form spontaneously, suggesting that the dimeric form may be the most efficient (or preferred) state of the complex in *Rba. sphaeroides*.

The small size and extremely high hydrophobicity of *Rba. sphaeroides* protein-U likely hindered its early discovery by biochemical and structural analyses. The presence of an additional protein in dimers of the *Rba. sphaeroides* LH1-RC was previously suspected from extra residual densities in an 8.5-Å cryo-EM projection map of 2D crystals from *Rba. sphaeroides* strain DD13/DGa2⁹. Two significant density features were identified within each half of the LH1-RC dimer, corresponding approximately to the positions of protein-U and PufX in our cryo-EM structure of the monomeric LH1-RC. One of the densities was tentatively proposed to be PufX⁹ but this corresponds to protein-U in our structure, while another relatively diffuse density feature was left unassigned⁹ and this corresponds to PufX in our study. Subsequently, a 3D structure of the dimeric LH1-RC from *Rba. sphaeroides* strain DBCΩG was determined at 8 Å using combined techniques²³. However, PufX was found at yet a different position in this structure (Supplementary Fig. 5). The two *Rba. sphaeroides* strains used in these studies both lacked a peripheral light-harvesting complex (LH2) but were mutant derivatives of the same parental strain (wild-type: NCIB8253)³⁸. Thus, the conflicting results likely indicate the difficulty in obtaining a definitive structure of such a large photocomplex at limiting resolutions, although structural rearrangements induced by genetic manipulations and/or dimerization cannot be absolutely ruled out.

Structures for individual polypeptides of LH1-β and PufX isolated from *Rba. sphaeroides* have been investigated using solution NMR. The β-polypeptide shows a single transmembrane α-helix starting from Leu19 with a remarkably high degree of disorder for the N-terminal domain in organic solvents³⁹. The flexible and unstructured features for the N-terminal region of the β-polypeptides were also observed in our cryo-EM structure in which the first 5–10 residues were invisible in the density map. Subsequently, the solution structures of the β-polypeptide were further determined in organic and detergent solvents, both exhibiting two helical domains separated by a flexible region with largely bent conformations^{40, 41} in contrast to that observed in our cryo-EM structure. Solution structures of PufX were also determined in organic solvents, yielding two quite different conformations: one showed two hydrophobic helices connected by a helical bend (PDB: 2ITA, 2NRG; Supplementary Fig. 10b)^{42, 43}, whereas a second revealed a rather straight, single continuous helix for the central transmembrane domain (PDB: 2DW3)⁴⁴. The latter closely matches that in our cryo-EM structure over the α-helical region (Supplementary Fig. 10a).

Alignments of eighteen PufX and PufX-like sequences in the database revealed a well-conserved central helical domain flanked by flexible termini (Supplementary Fig. 10c), and the diversities in structure, function and interactions of PufX have been thoroughly reviewed¹⁹. Genetic studies have demonstrated that *Rba. sphaeroides* PufX is most tightly associated with the LH1-RC complex, less tightly bound to the LH1-only complex, and least tightly associated with RC-only preparations⁴⁵. PufX tends to co-purify with the LH1 α-polypeptide⁴⁶, implying preferential interactions with the LH1 inner ring, consistent with our cryo-EM structure. Cross-species studies of PufX led to a hypothesis that in all *Rhodobacter* species PufX defines the composition of the monomeric LH1-RC complex, managing the aggregation state of the pigment proteins around the RC and determining the position of the LH1-ring opening relative to the entrance of the Q_B pocket in the RC¹². The nearly symmetric arrangements of PufX and protein-U relative to the LH1-ring opening and the fact that

deletion of protein-U resulted in a greatly reduced amount of dimeric LH1-RC may suggest a similar and/or cooperative role for protein-U in controlling the topology of the LH1 complex.

PufX in *Rba. sphaeroides* likely correlates with the carotenoid content in LH1. It was demonstrated that an engineered *Rba. sphaeroides* LH1 with a point-mutation (LH1- α Trp8 \rightarrow Phe) produced fewer carotenoids and allowed photosynthetic growth in the absence of PufX⁴⁷. This result was interpreted by assuming that an LH1 with fewer carotenoids would possess multiple pores (channels) in the closed complex as seen in the LH1 of *Tch. tepidum*²⁶ and that this would allow quinone transport through the pores in the absence of PufX and support photosynthetic growth. By contrast, a corresponding *Rba. sphaeroides* strain with a higher carotenoid content in its LH1 was photosynthetically incompetent⁴⁷. This in turn suggested that the high carotenoid content (approximately two carotenoids per $\alpha\beta$ -pair) in the native *Rba. sphaeroides* LH1 could block the LH1 channels (as confirmed in our structure, Fig. 4d) requiring alternative pathways for quinone transport in order to grow photosynthetically. Consequently, a PufX-containing C-shaped (monomeric) or S-shaped (dimeric) LH1 complex could fulfill this physiological requirement.

As experimental model system, phototrophic purple bacteria such as *Rba. sphaeroides* continue to reveal the most intimate secrets of photosynthetic energy conversion, and in this regard, our work provides the most robust and detailed structure of the *Rba. sphaeroides* LH1-RC structure to date. Our findings lay the foundation for an even deeper understanding of photosynthesis, in particular those aspects surrounding structure–function relationships in light-energy conversion, quinone transport, and regulation of carotenoid biosynthesis.

Note added

Just before we submit this manuscript to *bioRxiv*, a structure of LH1-RC-PufX complex from *Rba. veldkampii* has been published⁴⁸. Its LH1 complex contains 15 pairs of $\alpha\beta$ -polypeptides, 30 BChl *a* and 14 spheroidenes. The *Rba. veldkampii* core complex has no protein-U, but a PufX that is located at the same position with similar conformation to that of the PufX in the *Rba. sphaeroides* core complex of our work.

REFERENCES

1. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., and Rees, D. C. (1987) Structure of the reaction center from *Rhodobacter sphaeroides* R-26: The cofactors. *Proc. Natl. Acad. Sci. U.S.A.* 84, 5730-5734.
2. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., and Rees, D. C. (1987) Structure of the reaction center from *Rhodobacter sphaeroides* R-26: The protein subunits. *Proc. Natl. Acad. Sci. U.S.A.* 84, 6162-6166.
3. Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1985) Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3Å resolution. *Nature* 318, 618-624.
4. Olson, T. L., Williams, J. C., and Allen, J. F. (2014) The three-dimensional structures of bacterial reaction centers. *Photosynth. Res.* 120, 87-98.
5. Jungas, C., Ranck, J.-L., Rigaud, J.-L., Joliot, P., and Vermeglio, A. (1999) Supramolecular organization of the photosynthetic apparatus of *Rhodobacter sphaeroides*. *EMBO J.* 18, 534-542.

6. Bahatyrova, S., Frese, R. N., Siebert, C. A., Olsen, J. D., van der Werf, K. O., van Grondelle, R., Niederman, R. A., Bullough, P. A., Otto, C., and Hunter, C. N. (2004) The native architecture of a photosynthetic membrane. *Nature* 430, 1058-1062.
7. Scheuring, S., Francia, F., Busselez, J., Melandris, B. A., Rigaud, J.-L., and Levy, D. (2004) Structural role of PufX in the dimerization of the photosynthetic core complex of *Rhodobacter sphaeroides*. *J. Biol. Chem.* 279, 3620-3626.
8. Siebert, C. A., Qian, P., Fotiadis, D., Engel, A., Hunter, C. N., and Bullough, P. A. (2004) Molecular architecture of photosynthetic membranes in *Rhodobacter sphaeroides*: the role of PufX. *EMBO J.* 23, 690-700.
9. Qian, P., Hunter, C. N., and Bullough, P. A. (2005) The 8.5 Å projection structure of the core RC-LH1-PufX dimer of *Rhodobacter sphaeroides*. *J. Mol. Biol.* 349, 948-960.
10. Scheuring, S., Busselez, J., and Levy, D. (2005) Structure of the dimeric PufX-containing core complex of *Rhodobacter blasticus* by *in situ* atomic force microscopy. *J. Biol. Chem.* 280, 1426-1431.
11. Semchonok, D. A., Chauvin, J.-P., Frese, R. N., Jungas, C., and Boekema, E. J. (2012) Structure of the dimeric RC-LH1-PufX complex from *Rhodobaca bogoriensis* investigated by electron microscopy. *Phil. Trans. R. Soc. B* 367, 3412-3419.
12. Crouch, L. I., and Jones, M. R. (2012) Cross-species investigation of the function of the *Rhodobacter* PufX polypeptide and the composition of the RC-LH1 core complex. *Biochim. Biophys. Acta – Bioenergetics* 1817, 336-352.
13. Francia, F., Wang, J., Venturoli, G., Melandri, B. A., Barz, W. P., and Oesterhelt, D. (1999) The reaction center-LH1 antenna complex of *Rhodobacter sphaeroides* contains one PufX molecule which is involved in dimerization of this complex. *Biochemistry* 38, 6834-6845.
14. Francia, F., Wang, J., Zischka, H., Venturoli, G., and Oesterhelt, D. (2002) Role of the N- and C-terminal regions of the PufX protein in the structural organization of the photosynthetic core complex of *Rhodobacter sphaeroides*. *Eur. J. Biochem.* 269, 1877-1885.
15. Tsukatani, Y., Matsuura, K., Masuda, S., Shimada, K., Hiraishi, A., and Nagashima, K. V. P. (2004) Phylogenetic distribution of unusual triheme to tetraheme cytochrome subunit in the reaction center complex of purple bacteria. *Photosynth. Res.* 79, 83-91.
16. Klug, G., and Cohen, S. N. (1988) Pleiotropic effects of localized *Rhodobacter capsulatus* *puf* operon deletions on production of light-absorbing pigment-protein complexes. *J. Bacteriol.* 170, 5814-5821.
17. Farchaus, J. W., Barz, W. P., Grünberg, H., and Oesterhelt, D. (1992) Studies on the PufX polypeptide and its requirement for photoheterotrophic growth in *Rhodobacter sphaeroides*. *EMBO J.* 11, 2779-2788.
18. Lilburn, T. G., Haith, C. E., Prince, R. C., and Beatty, J. T. (1992) Pleiotropic effects of *pufX* gene deletion on the structure and function of the photosynthetic apparatus of *Rhodobacter capsulatus*. *Biochim. Biophys. Acta – Bioenergetics* 1100, 160-170.
19. Holden-Dye, K., Crouch, L. I., and Jones, M. R. (2008) Structure, function and interaction of the PufX protein. *Biochim. Biophys. Acta – Bioenergetics* 1777, 613-630.
20. Busselez, J., Cottevieille, M., Cuniasse, P., Gubellini, F., Boisset, N., and Lévy, D. (2007) Structural basis for the PufX-mediated dimerization of bacterial photosynthetic core complex. *Structure* 15, 1674-1683.
21. Liu, L.-N., Sturgis, J. N., and Scheuring, S. (2011) Native architecture of the photosynthetic membrane from *Rhodobacter veldkampii*. *J. Struct. Biol.* 173, 138-145.
22. Broglie, R. M., Hunter, C. N., Delepelaire, P., Niederman, R. A., Chua, N.-H., and Clayton, R. K. (1980) Isolation and characterization of the pigment-protein complexes

- of *Rhodopseudomonas sphaeroides* by lithium dodecyl sulfate/polyacrylamide gel electrophoresis. *Proc. Natl. Acad. Sci. USA* 77, 87-91.
23. Qian, P., Papiz, M. Z., Jackson, P. J., Brindley, A. A., Ng, I. W., Olsen, J. D., Dickman, M. J., Bullough, P. A., and Hunter, C. N. (2013) Three-dimensional structure of the *Rhodobacter sphaeroides* RC-LH1-PufX complex: Dimerization and quinone channels promoted by PufX. *Biochemistry* 52, 7575-7585.
 24. Nagashima, K. V. P., Sasaki, M., Hashimoto, K., Takaichi, S., Nagashima, S., Yu, L.-J., Abe, Y., Gotou, K., Kawakami, T., Takenouchi, T., Shibuya, Y., Yamaguchi, A., Ohno, T., Shen, J.-R., Inoue, K., Madigan, M. T., Kimura, Y., and Wang-Otomo, Z.-Y. (2017) Probing structure–function relationships in early events in photosynthesis using a chimeric photocomplex. *Proc. Natl. Acad. Sci. U.S.A.* 114, 10906-10911.
 25. Satoh, T., Hoshino, Y., and Kitamura, H. (1976) *Rhodopseudomonas sphaeroides* forma sp. *denitrificans*, a denitrifying strain as a subspecies of *Rhodopseudomonas sphaeroides*. *Arch. Microbiol.* 108, 265-269.
 26. Yu, L.-J., Suga, M., Wang-Otomo, Z.-Y., and Shen, J.-R. (2018) Structure of photosynthetic LH1-RC supercomplex at 1.9 Å resolution. *Nature* 556, 209-213.
 27. Swainsbury, D. J. K., Qian, P., Jackson, P. J., Faries, K. M., Niedzwiedzki, D. M., Martin, E. C., Farmer, D. A., Malone, L. A., Thompson, R. F., Ranson, N. A., Canniffe, D. P., Dickman, M. J., Holten, D., Kirmaier, C., Hitchcock, A., and Hunter, C. N. (2021) Structures of *Rhodopseudomonas palustris* RC-LH1 complexes with open or closed quinone channels. *Sci. Adv.* 7, eabe2631.
 28. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567-580.
 29. Parkes-Loach, P. S., Law, C. J., Recchia, P. A., Kehoe, J., Nehrlich, S., Chen, J., and Loach, P. A. (2001) Role of the core region of the PufX protein in inhibition of reconstitution of the core light-harvesting complexes of *Rhodobacter sphaeroides* and *Rhodobacter capsulatus*. *Biochemistry* 40, 5593-5601.
 30. Qian, P., Martin, E. C., Ng, I. W., and Hunter, C. N. (2017) The C-terminus of PufX plays a key role in dimerisation and assembly of the reaction center light-harvesting 1 complex from *Rhodobacter sphaeroides*. *Biochim. Biophys. Acta – Bioenergetics* 1858, 705-803.
 31. Qian, P., Siebert, C. A., Wang, P., Canniffe, D. P., and Hunter, C. N. (2018) Cryo-EM structure of the *Blastochloris viridis* LH1-RC complex at 2.9 Å. *Nature* 556, 203-208.
 32. Tani, K., Kanno, R., Makino, Y., Hall, M., Takenouchi, M., Imanishi, M., Yu, L.-J., Overmann, J., Madigan, M. T., Kimura, Y., Mizoguchi, A., Humbel, B. M., and Wang-Otomo, Z.-Y. (2020) Cryo-EM structure of a Ca²⁺-bound photosynthetic LH1-RC complex containing multiple $\alpha\beta$ -polypeptides. *Nat. Commun.* 11, 4955.
 33. Xin, Y., Shi, Y., Niu, T., Wang, Q., Niu, W., Huang, X., Ding, W., Yang, L., Blankenship, R. E., Xu, X., and Sun, F. (2018) Cryo-EM structure of the RC-LH core complex from an early branching photosynthetic prokaryote. *Nature Commun.* 9, 1568.
 34. Tani, K., Kanno, R., Ji, X.-C., Hall, M., Yu, L.-J., Kimura, Y., Madigan, M. T., Mizoguchi, A., Humbel, B. M., and Wang-Otomo, Z.-Y. (2021) Cryo-EM structure of the photosynthetic LH1-RC complex from *Rhodospirillum rubrum*. *bioRxiv*, doi: 10.1101/2021.1105.1130.446358.
 35. Javadpour, M. M., Eilers, M., Groesbeek, M., and Smith, S. O. (1999) Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophys. J.* 77, 1609-1618.

36. Eilers, M., Patel, A. B., Liu, W., and Smith, S. O. (2002) Comparison of helix interactions in membrane and soluble α -bundle proteins. *Biophys. J.* 82, 2720-2736.
37. Dong, H., Sharma, M., Zhou, H.-X., and Cross, T. A. (2012) Glycines: role in α -helical membrane protein structures and a potential indicator of native conformation. *Biochemistry* 51, 4779-4789.
38. Jones, M. R., Fowler, G. J. S., Gibson, L. C. D., Grief, G. G., Olsen, J. D., Crielgaard, W., and Hunter, C. N. (1992) Mutants of *Rhodobacter sphaeroides* lacking one or more pigment-protein complexes and complementation with reaction centre, LH1, and LH2 genes. *Mol. Microbiol.* 6, 1173-1184.
39. Kikuchi, J., Asakura, T., Loach, P. A., Parkes-Loach, P. S., Shimada, K., Hunter, C. N., Conroy, M. J., and Williamson, M. P. (1999) A light-harvesting antenna protein retains its folded conformation in the absence of protein-lipid and protein-pigment interactions. *Biopolymers* 49, 361-372.
40. Conroy, M. J., Westerhuis, W. H. J., Parkes-Loach, P. S., Loach, P. A., Hunter, C. N., and Williamson, M. P. (2000) The solution structure of *Rhodobacter sphaeroides* LH1b reveals two helical domains separated by a more flexible region: structural consequences for the LH1 complex. *J. Mol. Biol.* 298, 83-94.
41. Sorgen, P. L., Cahill, S. M., Krueger-Koplin, R. D., Krueger-Koplin, S. T., Schenck, C. C., and Girvin, M. E. (2002) Structure of the *Rhodobacter sphaeroides* light-harvesting 1 b subunit in detergent micelles. *Biochemistry* 41, 31-41.
42. Tunnicliffe, R. B., Ratcliffe, E. C., Hunter, C. N., and Williamson, M. P. (2006) The solution structure of the PufX polypeptide from *Rhodobacter sphaeroides*. *FEBS Lett.* 580, 6967-6971.
43. Ratcliffe, E. C., Tunnicliffe, R. B., Ng, I. W., Adams, P. G., Qian, P., Holden-Dye, K., Jones, M. R., Williamson, M. P., and Hunter, C. N. (2011) Experimental evidence that the membrane-spanning helix of PufX adopts a bent conformation that facilitate dimerisation of the *Rhodobacter sphaeroides* RC-LH1 complex through N-terminal interactions. *Biochim. Biophys. Acta – Bioenergetics* 1807, 95-107.
44. Wang, Z.-Y., Suzuki, H., Kobayashi, M., and Nozawa, T. (2007) Solution structure of the *Rhodobacter sphaeroides* PufX membrane protein: Implications for the quinone exchange and protein-protein interactions. *Biochemistry* 46, 3635-3642.
45. Pugh, R. J., McGlynn, P., Jones, M. R., and Hunter, C. N. (1998) The LH1-RC core complex of *Rhodobacter sphaeroides*: interaction between components, time-dependent assembly, and topology of the PufX protein. *Biochim. Biophys. Acta – Bioenergetics* 1366, 301-316.
46. Recchia, P. A., Davis, C. M., Lilburn, T. G., Beatty, J. T., Parkes-Loach, P. S., Hunter, C. N., and Loach, P. A. (1998) Isolation of the PufX protein from *Rhodobacter capsulatus* and *Rhodobacter sphaeroides*: Evidence for its interaction with the α -polypeptide of the core light-harvesting complex. *Biochemistry* 37, 11055-11063.
47. Olsen, J. D., Martin, E. C., and Hunter, C. N. (2017) The PufX quinone channel enables the light-harvesting 1 antenna to bind more carotenoids for light collection and photoprotection. *FEBS Lett.* 591, 573-580.
48. Bracun, L., Yamagata, A., Christianson, B. M., Terada, T., Canniffe, D. P., Shirouzu, M., and Liu, L.-N. (2021) Cryo-EM structure of the photosynthetic RC-LH1-PufX supercomplex at 2.8-Å resolution. *Sci. Adv.* 7, eabf8864.

ACKNOWLEDGMENT

We thank Megumi Kobayashi, Ayumi Imai and Risa Onuma for providing excellent technical assistances. This research was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Numbers JP20am0101118 (support number 1758) and JP20am0101116 (support number 1878), 17am0101116j0001, 18am0101116j0002 and 19am0101116j0003. R.K., M.H. and B.M.H. acknowledge the generous support of the Okinawa Institute of Science and Technology and the Japanese Cabinet Office. L.-J.Y. acknowledges support of the National Key R&D Program of China (No. 2019YFA0904600). This work was supported in part by JSPS KAKENHI Grant Numbers JP16H04174, JP18H05153, JP20H05086 and JP20H02856, Takeda Science Foundation, and the Kurata Memorial Hitachi Science and Technology Foundation, Japan.

Author Contributions

Z.-Y.W.-O. and K.T. designed the work, K.T., K.V.P.N., R.K., S.K., R.K. and M.H. performed the experiments, K.T., K.V.P.N., R.K., L.-J.Y., Y.K., M.T.M., A.M., B.M.H. and Z.-Y.W.-O. analyzed data, Z.-Y.W.-O., K. T. and M.T.M. wrote the manuscript. All authors edited and revised the manuscript.

Competing interests: The authors declare no competing interests.

Data availability: Map and model have been deposited in the EMDB and PDB with the accession codes: EMD-31400 and PDB-7F0L. All other data are available from the authors upon reasonable request.