

## Asymmetric Structure of the Native *Rhodobacter sphaeroides* Dimeric LH1-RC Complex

K. Tani<sup>1,8\*</sup>, R. Kanno<sup>2,8</sup>, R. Kikuchi<sup>3</sup>, S. Kawamura<sup>3</sup>, K. V. P. Nagashima<sup>4</sup>, M. Hall<sup>2</sup>, A. Takahashi<sup>2</sup>, L.-J. Yu<sup>5</sup>, Y. Kimura<sup>6</sup>, M. T. Madigan<sup>7</sup>, A. Mizoguchi<sup>1</sup>, B. M. Humbel<sup>2</sup> & Z.-Y. Wang-Otomo<sup>3\*</sup>

<sup>1</sup> Graduate School of Medicine, Mie University, Tsu 514-8507, Japan

<sup>2</sup> 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

<sup>3</sup> Faculty of Science, Ibaraki University, Mito 310-8512, Japan

<sup>4</sup> Research Institute for Integrated Science, Kanagawa University, 2946 Tsuchiya, Hiratsuka, Kanagawa 259-1293, Japan

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

<sup>6</sup> Department of Agrobioscience, Graduate School of Agriculture, Kobe University, Nada, Kobe 657-8501, Japan

<sup>7</sup> School of Biological Sciences, Department of Microbiology, Southern Illinois University, Carbondale, IL 62901 USA

<sup>8</sup> These authors contributed equally to this work.

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

**Abstract:** The light-harvesting-reaction center (LH1-RC) core complex of purple photosynthetic bacterium *Rhodobacter (Rba.) sphaeroides* is characterized by the presence of both a dimeric form and a monomeric form. Following structure determination of the monomeric LH1-RC including its previously unrecognized component designated protein-U (*Nat. Commun.* **12**, 6300, 2021), here we present cryo-EM structures of the dimeric LH1-RC from native *Rba. sphaeroides* IL106 at 2.75 Å resolution and from an LH1-RC monomer lacking protein-U ( $\Delta U$ ) at 2.64 Å resolution. The native dimeric core complex reveals many asymmetric features in the arrangement of its two monomeric components including the structural integrity of protein-U, the overall LH1 organization, and the rigidities of the proteins and pigments that form the complex. PufX polypeptides play a critical role in connecting two monomers, with one PufX interacting at its N-terminus with another PufX and an LH1  $\beta$ -polypeptide in another monomer, in good agreement with biochemical analyses. One of the proteins-U was only partially identified in the dimeric structure, signaling significantly different degrees of disorder in the two monomers. The  $\Delta U$  LH1-RC monomer revealed a half-moon-shaped structure containing 11  $\alpha$ - and 10  $\beta$ -polypeptides (compared with 14 of each in the wild type), indicating a critical role for protein-U in controlling the number of  $\alpha\beta$ -subunits required for correct assembly and stabilization of the LH1-RC dimer. The structural features are discussed in relation to the unusual topology of intracytoplasmic photosynthetic membranes and an assembly model proposed for the native *Rba. sphaeroides* dimeric LH1-RC complex in membranes of wild-type cells.

## Introduction

*Rhodobacter (Rba.) sphaeroides* is the most thoroughly investigated purple phototrophic bacterium and is widely used as a model organism. The combination of its metabolic diversity, ease of growth and well-established genetics have revealed the details of many fundamental mechanisms of photosynthesis including photochemistry, metabolism, and regulation. The core light-harvesting (LH1) system of *Rba. sphaeroides* is characterized by the co-existence of dimeric and monomeric forms corresponding to the S- and C-shaped structures, respectively, observed in native membranes<sup>1,2</sup>. Both dimeric and monomeric *Rba. sphaeroides* LH1 complexes can be separately purified in a reaction center (RC)-associated form. The so-called core complex (LH1-RC) contains a protein, PufX<sup>3,4</sup>, whose roles in regulation of membrane morphology, core photocomplex organization, and cyclic electron transfer have been intensively investigated<sup>5</sup> but still unsettled.

In a recent study, we determined the structure of a monomeric LH1-RC from *Rba. sphaeroides* strain IL106<sup>6,7</sup>. The LH1 complex forms a C-shaped structure composed of 14  $\alpha\beta$ -polypeptides around the RC with a large ring opening. In addition to PufX, a previously unrecognized integral membrane protein, referred to as protein-U, was identified in the core complex. Protein-U is located opposite PufX on the LH1-ring opening and has a U-shaped conformation. Deletion of protein-U resulted in a mutant strain that expressed a much-reduced amount of dimeric LH1-RC, implying key roles for protein-U in stabilization of the dimeric form of LH1-RC complex<sup>7</sup>. A similar structure was also reported for a monomeric core complex derived from a genetically modified strain of *Rba. sphaeroides*<sup>8</sup>.

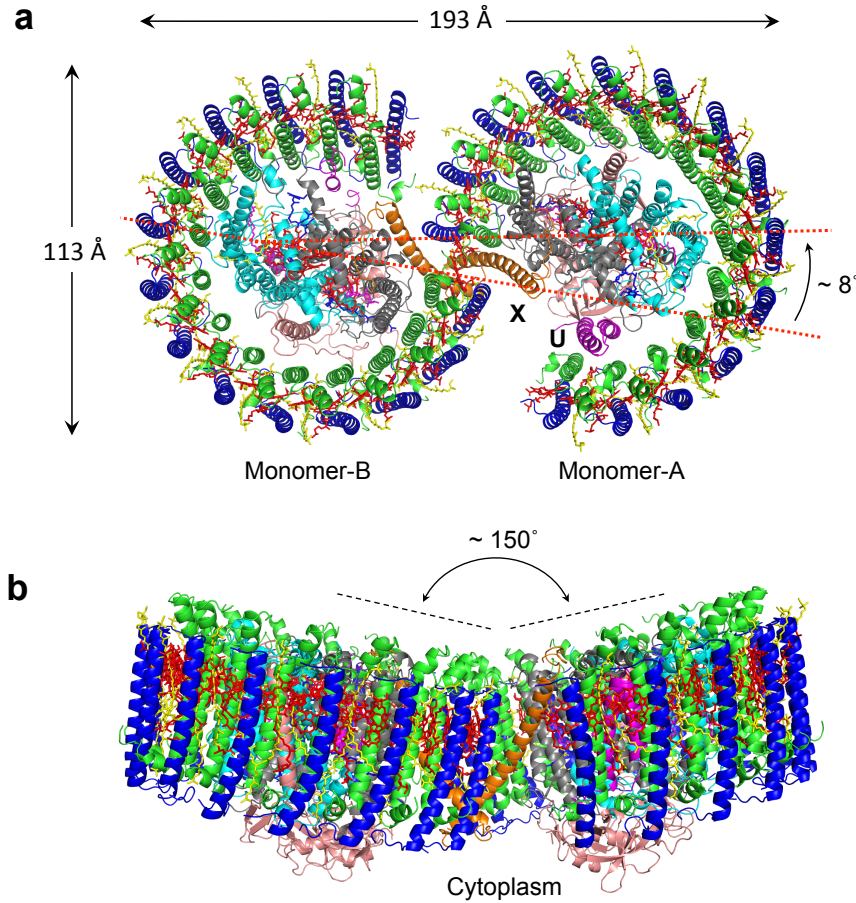
Much effort has been devoted to determining the structure of the dimeric LH1-RC from *Rba. sphaeroides*<sup>8-12</sup>. Recently, a two-fold symmetric structure of the dimeric LH1-RC was reported for an LH2-deficient mutant strain of *Rba. sphaeroides* (strain DBC $\Omega$ G)<sup>8</sup>. Dimeric LH1-RC complexes were also observed in *Rba. blasticus*<sup>13</sup> and *Rhodobaca bogoriensis*<sup>14</sup>, and were suggested to exist in a few other purple bacteria<sup>15</sup>. Here, we present cryo-EM structures of the dimeric core complex from native *Rba. sphaeroides* IL106 and the monomeric LH1-RC from a protein-U-deleted mutant strain (strain  $\Delta$ U) of *Rba. sphaeroides* IL106. In contrast to the symmetric structure reported from the LH2-deficient mutant strain<sup>8</sup>, the native LH1-RC dimer revealed an asymmetric S-shaped structure. Based on our structural information, the functions of PufX and protein-U in the *Rba. sphaeroides* LH1-RC core complex have become more clear and are discussed in the context of a model that describes the assembly and stabilization of the mature LH1-RC complex as it exists in wild-type cells.

## Results

**Asymmetric Features of the Native *Rba. sphaeroides* LH1-RC Dimer.** The cryo-EM structure of the dimeric LH1-RC complex from native *Rba. sphaeroides* IL106 was determined at 2.75 Å resolution with C1 symmetry (Fig. 1, Supplementary Table 1 and Supplementary Figs. 1–5). The two halves in the LH1-RC dimer, designated as monomer-A and monomer-B, are not symmetric. While two PufX polypeptides and all LH1 and RC polypeptides were fully identified in the dimer, protein-U in monomer-B was only partially detected, presumably due to its large disorder. No other proteins were detected in the native dimer structure.

Monomer-A is rotated approximately 8° anticlockwise relative to the mirror image of monomer-B as determined from the dihedral angle of cross sections in the two monomers that are perpendicular to the presumed membrane surface (Fig. 1a). This rotation results in a significantly larger opening in the LH1 ring of monomer-A than the LH1 ring of monomer-B. The two monomers in the dimeric LH1-RC are arranged in a bent configuration with an angle of approximately 150° inclined toward each other of their presumed periplasmic

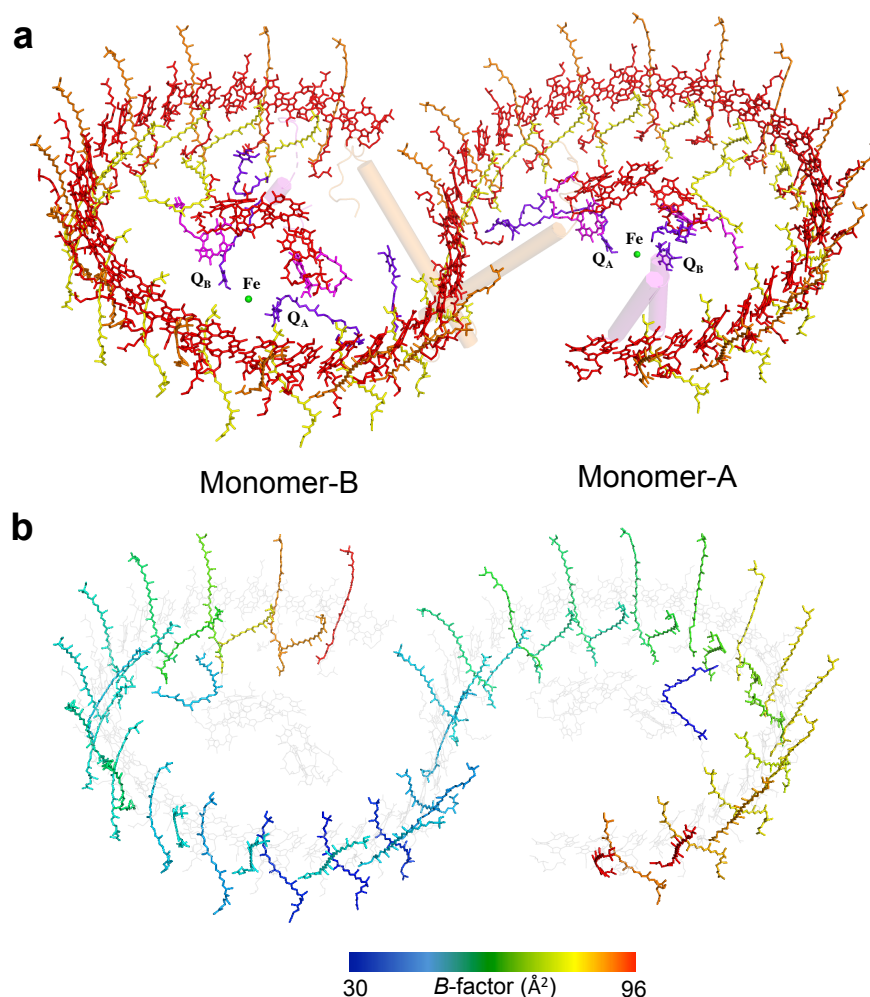
surfaces (Fig. 1b). This topology is similar to that of a dimeric LH1-RC structure from the LH2-deficient *Rba. sphaeroides* mutant strain DBCΩG<sup>8,12</sup>, and is consistent with the curvature of vesicle-type intracytoplasmic membrane (ICM) surface as observed in phototrophically grown native *Rba. sphaeroides* cells<sup>16</sup>.



**Figure 1. Structure overview of the dimeric LH1-RC complex from *Rba. sphaeroides* IL106.** (a) Top view from periplasmic side of the membrane with PufX (orange) and protein-U (magenta) in monomer-A indicated by their chain ID: X and U, respectively. (b) Side view of the core complex parallel to the membrane plane with the H-subunits (salmon) at the bottom. Color scheme: LH1- $\alpha$ , green; LH1- $\beta$ , blue; RC-L, black; RC-M, cyan; BChl *a*, red sticks; spheroidenes, yellow sticks; BPhe *a*, magenta sticks.

The *Rba. sphaeroides* LH1 dimeric complex is composed of 28 pairs of  $\alpha\beta$ -polypeptides, 56 bacteriochlorophylls (BChl) *a* and 52 spheroidenes surrounding two RCs with an S-shaped structure. BChl *a* molecules in the *Rba. sphaeroides* dimeric LH1 form a partially overlapping S-shaped ring with average Mg–Mg distances of 9.5 Å within an  $\alpha\beta$ -subunit and 8.3 Å between the  $\alpha\beta$ -subunits (Fig. 2a). These values are similar to those observed for the monomeric *Rba. sphaeroides* LH1 (Supplementary Table 2)<sup>7</sup>. The Q<sub>A</sub> site of quinones in the RC is located close to PufX, whereas the Q<sub>B</sub> site is on the side of protein-U. As in the case of monomeric *Rba. sphaeroides* LH1-RC<sup>7</sup>, two groups of all-*trans*-spheroidenes with distinct configurations were detected in the dimeric LH1 (Fig. 2b). Again, asymmetric features were observed for the B-factor (an indicator of the degree of disorder) of these carotenoids in LH1 (Fig. 2b) as follows: (i) the carotenoids in monomer-A have larger

*B*-factors than those at corresponding positions in monomer-B; (ii) group-B carotenoids (defined as those protruding on the periplasmic surface)<sup>7</sup> have larger *B*-factors than those of the group-A carotenoids (defined as those embedded deeply in the transmembrane region) in an  $\alpha\beta$ -subunit; and (iii) the carotenoids around the dimer junction have smaller *B*-factors than do those of the carotenoids near the ends of the LH1-ring. These results indicate that the carotenoids in monomer-B are generally more rigid than those in monomer-A, and the carotenoids in the ring junction are more rigid than carotenoids in other portions of the LH1 ring with similar rigidity to the *cis*-spheroidenes in the RCs. These trends are also consistent with the resolution and *B*-factor distributions over all atoms (Supplementary Fig. 3).

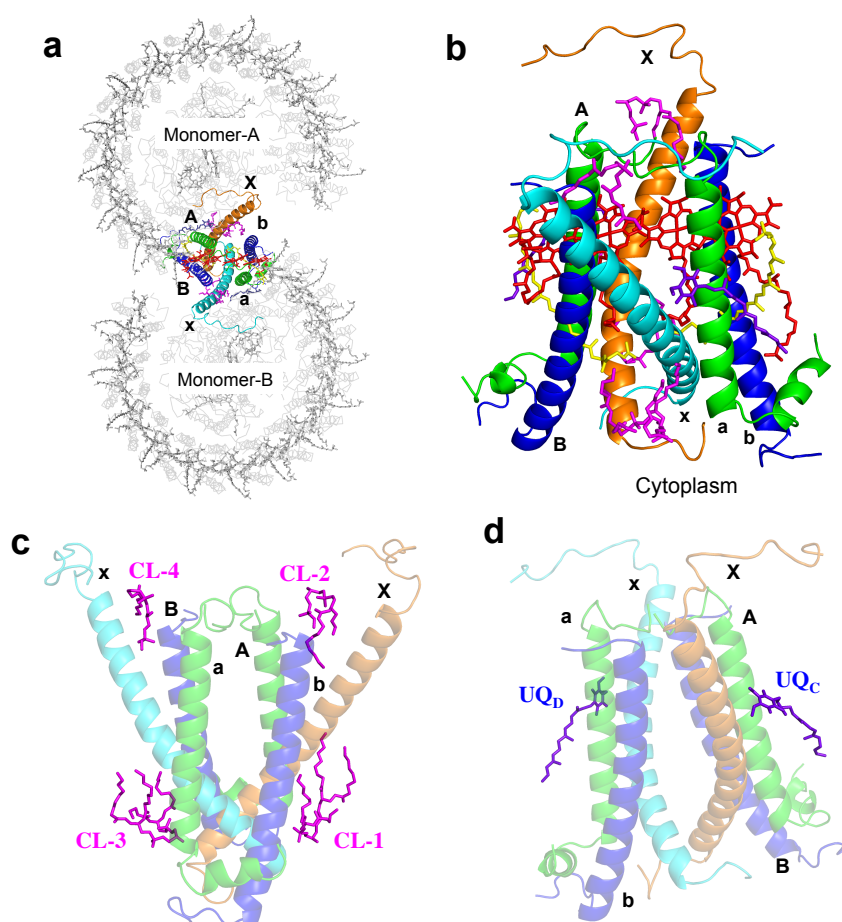


**Figure 2. Arrangement of the cofactors in the dimeric LH1-RC complex.** (a) Tilted view from the periplasmic side of the membrane with PufX (orange) and proteins-U (magenta) shown by transparent cylinders. Color scheme: BChl *a*, red sticks; spheroidenes (group-A), yellow sticks; spheroidenes (group-B), orange sticks; BPhe *a*, magenta sticks; UQ-10, purple sticks; non-heme Fe, green sphere. (b) *B*-factor distribution of the carotenoids in the core complex with colors representing the *B*-values indicated in the color bar. BChl *a* and BPhe *a* molecules are colored light-gray.

**Junction of the *Rba. sphaeroides* LH1-RC Dimer.** Monomers-A and -B in the dimeric LH1-RC join at the LH1-ring ends near the PufX (Fig. 3a). The junction of the two monomers is composed of two PufX and two pairs of LH1  $\alpha\beta$ -polypeptides (Fig. 3b). There are no carotenoids present at the joint (between chains AB and ab pairs) of the two rings

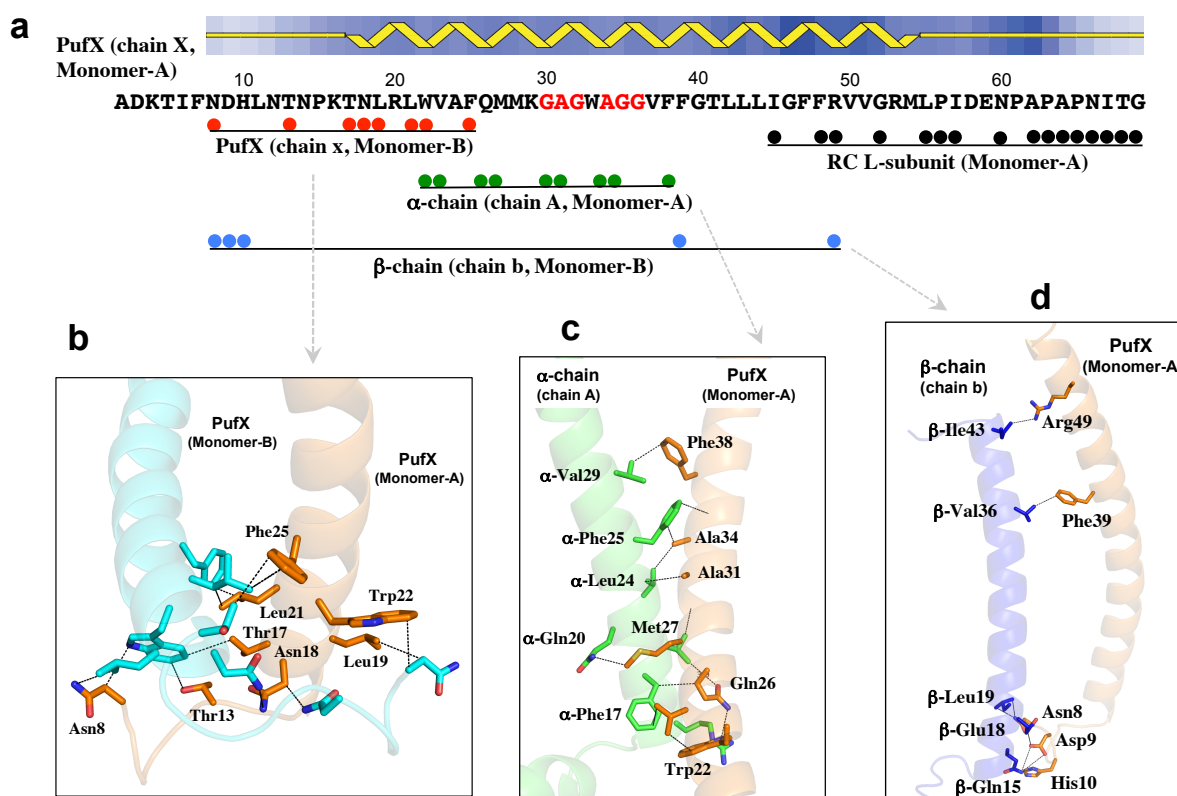


occupied by PufX (bottom panel in Supplementary Fig. 5). The two PufX polypeptides function as linkers connecting the two monomers (Fig. 4), and due to a largely tilted conformation, the two PufX polypeptides can only interact with each other through their N-terminal domains on the cytoplasmic side of the membrane (Fig. 4a, 4b). Interactions between charged/polar residues and CH- $\pi$  interactions primarily occur in the range of Asn8–Phe25. Four tightly bound cardiolipin (CL) and two ubiquinone (UQ) molecules were identified in the junction (Fig. 3 and Supplementary Fig. 6). The head groups of CL-1 and CL-3 interact extensively with the charged/polar residues in the N-terminal domains of the two PufX (Supplementary Fig. 6a), strengthening the connection between the two monomers. CL-2 and CL-4 also interact with PufX and are in close contacts with a different  $\beta$ -polypeptide on the periplasmic side of the membrane (Supplementary Fig. 6b). The head groups of two UQ molecules are in close proximity to PufX and LH1- $\alpha$  polypeptides in the transmembrane region (Supplementary Fig. 6c).



**Figure 3. Junction of the *Rba. sphaeroides* LH1-RC dimer.** (a) Top view from the periplasmic side of the membrane with the polypeptides and cofactors in the dimer junction shown in colors and by their chain ID: A, B and X denote the  $\alpha$  (green),  $\beta$  (blue) and PufX (orange) polypeptides in monomer-A, respectively; a, b and x denote the  $\alpha$  (green),  $\beta$  (blue) and PufX (cyan) polypeptides in monomer-B, respectively. BChl *a* and spheroidenes are shown by red and yellow sticks, respectively. Cardiolipin (CL) and UQ molecules are shown by magenta and purple sticks, respectively. (b) Side view of the expanded dimer junction with the same color scheme as in (a). (c) Four CL molecules (magenta sticks) in the dimer junction. (d) Two UQ molecules (purple sticks) in the dimer junction.

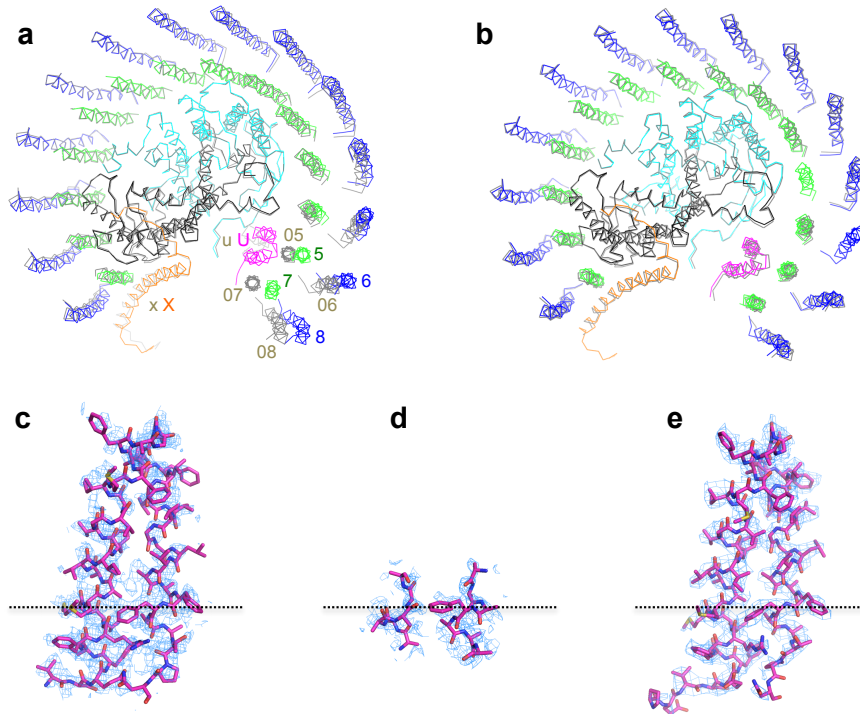
Extensive hydrophobic interactions were observed in the transmembrane region between PufX and the LH1  $\alpha$ -polypeptide positioned at the junction within the same monomer (Fig. 3, Fig. 4a, 4c). Moreover, several CH-p interactions along with cation-p interactions<sup>7</sup> exist between the two polypeptides. PufX crosses an  $\alpha$ -polypeptide through a short, extremely rigid Gly- and Ala-rich stretch (GAGWAGG) (Fig. 4a, Supplementary Fig. 7) as was demonstrated by H/D exchange NMR<sup>17</sup>. The center of the stretch in PufX points to the Gly21 in the  $\alpha$ -polypeptide that is conserved in all *Rhodobacter* species. As a result, the contact interface between the two polypeptides is composed of two crossing grooves formed by the residues with small sidechains (essentially Gly), allowing the closest possible approach ( $< 4.0$  Å) of the two main chains through van der Waals interactions. A similar feature was observed for a PufX/ $\alpha$ -polypeptide pair in the monomeric LH1-RC-PufX structure from *Rba. veldkampii*; in this structure, PufX crosses the  $\alpha$ -polypeptide through a narrow segment (Gly-Met-Gly) between PufX and the Gly21 in the  $\alpha$ -polypeptide (Supplementary Fig. 7c, 7d)<sup>18</sup>. It is notable that the Gly30 of PufX and Gly21 of the  $\alpha$ -polypeptide in *Rba. sphaeroides* are conserved in all *Rhodobacter* species. Other, larger residues around the crossing interface in *Rba. sphaeroides* PufX have most of their bulky sidechains pointing outward (Supplementary Fig. 7b).



**Figure 4. Interaction mapping of the PufX in monomer-A with its partner proteins.** (a) Primary sequence and secondary structure of the expressed PufX with the color spheres representing its amino acid residues that are in close contacts ( $< 4.0$  Å) with partner polypeptides as indicated. (b) Interactions between the two PufX polypeptides in their N-terminal regions with the dashed lines representing typical close contacts between the two polypeptides. (c) Close contacts ( $< 4.0$  Å) between PufX and an  $\alpha$ -polypeptide of the same monomer in their transmembrane regions. The crossing portion in PufX corresponds to a short, Gly- and Ala-rich stretch as indicated by red fonts in the sequence of (a). (d) Close contacts between PufX and a  $\beta$ -polypeptide of a different monomer in their N-terminal and transmembrane regions.

PufX also interacts with the  $\beta$ -polypeptide that is positioned at the junction and is part of a different monomer (Fig. 3, Fig. 4a, 4d). Interactions between charged/polar residues in the N-terminal domains (Asn8–His10) and CH- $\pi$  interactions in the transmembrane domains were observed between these two polypeptides. PufX interacts extensively in the C-terminal domain with the RC L-subunit of the same monomer (Fig. 4a). Such interactions have been described in detail for a monomeric *Rba. sphaeroides* LH1-RC<sup>7</sup> and an LH1-RC-PufX complex from *Rba. veldkampii*<sup>18</sup>. The structural features of PufX in our work are consistent with biochemical analyses showing that the PufX is required for dimerization of the core complex<sup>19-22</sup> and that truncations of N- and C-terminal domains of PufX also disrupted formation of a dimeric LH1-RC<sup>21,23</sup>.

**Structural Comparison of the Monomers in the Dimer.** Asymmetry was also observed for the LH1 organization of the *Rba. sphaeroides* dimeric LH1-RC structure (Supplementary Fig. 4c). The two monomers revealed different arrangements of the LH1 subunits, especially for the two  $\alpha\beta$ -pairs near protein-U (Fig. 5a). Superposition of the RC M-subunits between the two monomers showed that the 13–14th  $\alpha\beta$ -polypeptides in monomer-B have gradually shifted inward from their corresponding pairs in monomer-A at the end of the LH1 rings, reaching a maximum deviation of approximately 6 Å for the last 14th  $\alpha\beta$ -pair (chain ID: 07/08) (Fig. 5a). The inward shift of the last two  $\alpha$ -polypeptides (chain ID: 05 and 07) in monomer-B could cause a conformational disruption of the adjacent protein-U and explain the large disorder observed for protein-U in monomer-B.

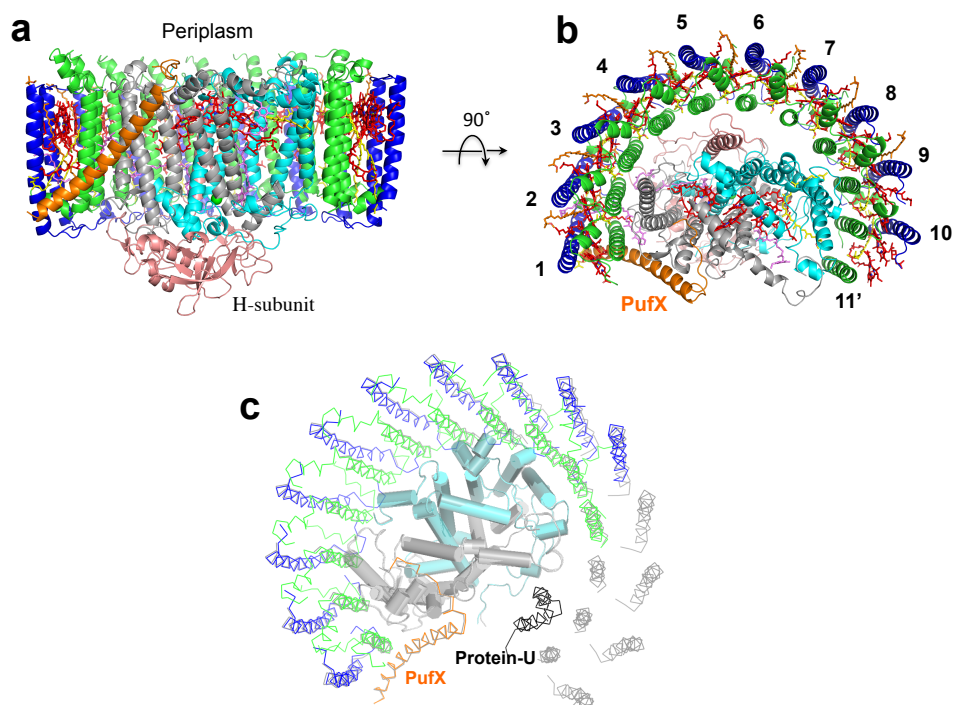


**Figure 5. Structural comparison of the monomers in the dimer.** (a) Top view of superposition of the C $\alpha$  carbons of the RC M-subunits between monomer-A (colored) and monomer-B (gray) shown by ribbon representations and viewed from periplasmic side. Color scheme for the proteins in monomer-A is the same as in Fig. 1. PufX (X, x) and proteins-U (U, u) are indicated by their chain ID (X and U in monomer-A; x and u in monomer-B). The RC H-subunits are omitted for clarity. Two  $\alpha\beta$ -subunits in the monomer-B indicated by their chain ID 05, 06, 07 and 08 are largely deviated from

the corresponding polypeptides in monomer-A (chain ID 5, 6, 7 and 8). **(b)** Similar superposition between monomer-A (colored) and a monomeric *Rba. sphaeroides* LH1-RC (gray, PDB: 7F0L) reported previously. **(c–e)** Structures of proteins-U with their density maps at a contour level of  $4\sigma$  in monomer-A **(c)**, monomer-B **(d)**, and the monomeric core complex **(e)**. Dotted lines indicate the position of Phe46 of protein-U. Most parts of the protein-U in monomer-B are disordered.

The structure of monomer-A in the dimeric LH1-RC well overlaps with that of a monomeric *Rba. sphaeroides* LH1-RC<sup>7</sup> from the same wild-type strain of *Rba. sphaeroides* for all protein and pigment components (Fig. 5b). The two copies of protein-U in the dimeric core complex exhibited significantly different degrees of disorder (Fig. 5c–d). The protein-U in monomer-B could only be traced from fragmented densities for 13 amino acids near the cytoplasmic side (Fig. 5d), while other portions were highly disordered. This is likely due to the anticlockwise rotation of monomer-A toward the LH1-ring edge of monomer-B where protein-U in monomer-B is located (Fig. 1a). This rotation would naturally push the last two  $\alpha\beta$ -subunits in monomer-B toward the nearby protein-U (Fig. 5a) and as a result, disrupt its conformation.

**Structure of the *Rba. sphaeroides* LH1-RC  $\Delta$ U Monomer.** In order to further explore structure-function relationships of protein-U, we also determined the cryo-EM structure of a monomeric LH1-RC from a protein-U deletion strain of *Rba. sphaeroides* IL106 (strain  $\Delta$ U) at 2.64 Å resolution (Supplementary Fig. 8). The  $\Delta$ U LH1-RC monomer revealed a half-moon-shaped structure containing 11  $\alpha$ - and 10  $\beta$ -polypeptides (Fig. 6). Three  $\alpha\beta$ -pairs near where protein-U resides in the wild-type LH1-RC monomer as well as the  $\beta$ -polypeptide paring with the 11th  $\alpha$ -polypeptide were not identified in the density map (Fig. 6c). In addition, a group-B spheroidene between the 9th and 10th  $\beta$ -polypeptides was also not observed (Fig. 6b). By contrast, all other proteins and pigments in the  $\Delta$ U LH1-RC were present and had the same structures as those in the wild-type monomeric LH1-RC<sup>7</sup>. This result clearly indicates that protein-U plays a critical role in controlling the number of  $\alpha\beta$ -subunits in an LH1 monomer and that a structurally complete LH1 ring is a prerequisite for both LH1-RC dimer formation and its subsequent stabilization, in firm agreement with biochemical analyses<sup>7</sup>.





**Figure 6. Structure of the *Rba. sphaeroides* LH1-RC  $\Delta$ U monomer.** (a) Side view of the complex with PufX (orange) in the front. Color scheme is same as in Fig. 1. (b) Top view from periplasmic side of the membrane with numbering for the LH1 polypeptides. (c) Overlap view of the  $\Delta$ U monomer (colored) and wild-type monomeric *Rba. sphaeroides* LH1-RC (gray, PDB: 7F0L) by superposition of C $\alpha$  carbons of the RC M-subunits. Protein-U as it exists in the wild-type monomer is shown by a black ribbon.

## Discussion

At the resolution of 2.75 Å achieved in our work, the dimeric structure of the native *Rba. sphaeroides* LH1-RC core complex revealed several asymmetric characteristics. These include the arrangement of the two constituting monomers, the structural integrity of protein-U, the organization of LH1, and the rigidity of protein and pigment components in the dimer. These stand in contrast to a two-fold symmetric dimer structure reported for the core complex from LH2-deficient *Rba. sphaeroides* strain DBCΩG<sup>8</sup>. The asymmetric features observed in our work may well be related to each other and may have their origins in the bent configuration of the dimer (Fig. 1b) that is likely imposed by the topology of ICM that house the photocomplexes of purple bacteria.

ICM is formed by invagination of cytoplasmic membrane (CM) with an inverted curvature relative to that of the CM and is known to accommodate major photosynthetic pigment-membrane protein machineries<sup>24</sup>. The ICMs in wild-type *Rba. sphaeroides* cells were demonstrated to have a spherical shape<sup>25,26</sup>. Upon isolation of ICM (referred to as chromatophores), the periplasmic surface of photosynthetic membranes is positioned on the curved-in side of the chromatophores (inside of the spherical vesicles)<sup>27</sup>, consistent with the curvature of a bent configuration of the LH1-RC dimer. The uneven rigidities over a dimension of approximately 190 Å of the native dimeric LH1-RC (Fig. 2b, Supplementary Fig. 3) likely reflect the inhomogeneity and fluidity of the photosynthetic membranes<sup>28</sup>, and the carotenoid molecules distributed in the dimer can be used as indicators of rigidity for the entire core complex.

In contrast to the structure of chromatophores from wild-type *Rba. sphaeroides*, ICM from a mutant strain lacking LH2 exhibited a flattened tubular morphology<sup>29</sup>. These differences may be the result of one (or both) of two possibilities. First, our own results have shown that ICM from cells lacking LH2 have a significantly different phospholipid composition than that of ICM from wild-type cells. Specifically, ICM from cells lacking LH2 has a remarkably lower content of phosphatidylglycerol and higher content of phosphatidylcholine than does ICM from wild-type cells<sup>30</sup>, and these differences, coupled with the total absence of LH2, likely change native ICM structure in dramatic ways. And secondly, a new protein has been reported from dimeric LH1-RC complexes prepared from the same LH2-deficient strain, a protein not observed in the dimeric complexes we prepared from wild-type cells. This protein, referred to as protein Z<sup>8</sup>, is also absent from monomeric LH1-RC complexes from the LH2-deficient mutant<sup>31</sup>; because of this specific association, protein Z was proposed to play a role in the formation of LH1-RC dimers<sup>8</sup>. On the contrary, our structure of the native complex yielded no evidence for a Z-like protein, it should not be necessary for formation of dimeric complexes in wild-type cells.

PufX in the *Rba. sphaeroides* LH1-RC complex plays a critical role in formation of the dimeric core complex, and our cryo-EM structure of PufX in a dimer provides a foundation for examining its structural and functional roles that have been investigated only biochemically thus far. Deletion of the gene encoding *Rba. sphaeroides* PufX results in a mutant strain that can only synthesize a monomeric form of the core complex<sup>19-22</sup>. Moreover, truncations of either the N- or C-terminal domain<sup>21,23</sup> or point-mutations of certain PufX

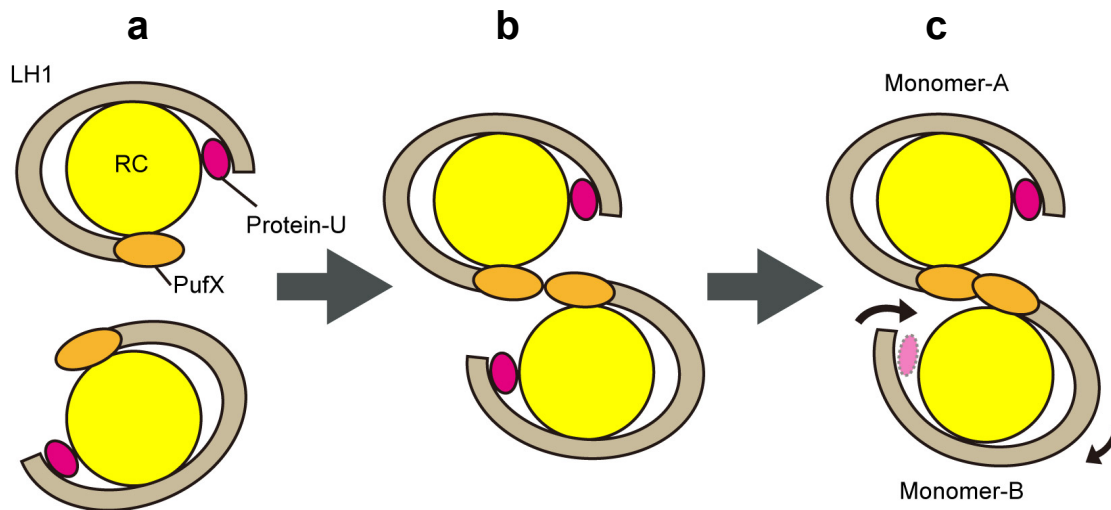
residues<sup>21,22</sup> also disabled dimer formation. These results now can be explained by an interaction map of PufX that shows that one LH1-RC monomer interacts with the PufX and a  $\beta$ -polypeptide of a second monomer primarily through their N-terminal domains (Asn8–Phe25 in PufX, Fig. 4), thus connecting the two monomers. This structure is consistent with mutagenic studies that show that removal of 12 or more residues from the N-terminus of PufX results in the complete loss of dimeric complexes even though the truncated PufX polypeptides are expressed and some of the mutants still retained the ability to grow photosynthetically<sup>21,23</sup>. Four CL molecules interacting with PufX and  $\beta$ -polypeptides in the junction also stabilize the dimeric form (Fig. 3c). However, these PufX-mediated inter-monomer interactions can be disrupted *in vitro* with relatively high concentrations of the detergent *n*-octyl- $\beta$ -glucopyranoside<sup>19</sup>, resulting exclusively in monomeric forms.

*Rba. sphaeroides* PufX interacts with the RC L-subunit in a monomer over a wide range of its C-terminal domain (Ile45–Gly69, Fig. 4a)<sup>7</sup>. Point-mutations of the PufX Arg49, Gly52 and Arg53 to Leu resulted in loss of the dimeric LH1-RC; nevertheless, these mutant strains still grow photosynthetically at rates similar to that of wild-type<sup>22</sup>. The Arg49 and Gly52 in PufX are in close contacts with the RC L-subunit near the periplasmic surface as shown in Fig. 4a. By contrast, deletions of 7 or more residues from the C-terminus that are encoded in the *pufX* gene result in the absence of the truncated PufX in core complexes and assembly of only monomeric forms<sup>23</sup>; such results point to a major role for the C-terminus of PufX in its assembly in the monomeric complex. Compared with N- and C-terminal domains, hydrophobic interactions in the transmembrane regions of PufX and an  $\alpha$ -polypeptide of the same monomer is likely less critical for dimer formation. For example, mutations of several Gly residues to Leu in the central stretch of PufX did not decrease the propensity of the core complex to assemble in a dimeric form<sup>20</sup>. Overall, the interaction features of PufX with its partner proteins revealed in our study are in good agreement with biochemical analyses and support the conclusion that the N-terminal domain of PufX is directly responsible for dimerization while the C-terminal domain mainly participates in assembly of the core complex and along with the transmembrane region stabilizes the dimeric structure.

Another important finding of our work concerns the role of protein-U. Although protein-U has no inter-monomeric interactions in the *Rba. sphaeroides* dimeric LH1-RC, it clearly plays a role in regulating LH1 ring size. Appropriate ring size is crucial for assembly and stabilization of the dimeric complex in order for it to function in a curved membrane topology. This is consistent with biochemical analyses that show that strain  $\Delta U$  can still produce a dimeric complex but at much reduced levels compared with the wild type<sup>7</sup>. A recent structure of the LH1-RC from *Rba. veldkampii* showed that its LH1 is composed of 15  $\alpha\beta$ -subunits (Supplementary Fig. 9d)<sup>18</sup>, one subunit more than in the *Rba. sphaeroides* LH1. *Rba. veldkampii* only synthesizes the monomeric form of LH1-RC and lacks protein-U in its core complex, further supporting our conclusion that protein-U in *Rba. sphaeroides* regulates the number of  $\alpha\beta$ -subunits in the complex, thereby controlling the size of the LH1-ring opening.

Based on the structures determined in our work and that of others, we propose a model for assembly of the dimeric LH1-RC complex in native membranes of wild-type *Rba. sphaeroides* (Fig. 7). In our model, the process begins when two structurally complete C-shaped LH1-RC monomers collide and link at their PufX sites. Protein-U does not directly contribute at this stage but ensures the correct LH1 ring size required for initial dimerization. The dimer retains a symmetric conformation (Fig. 7b) at this point but is likely in a transient state because of the instability associated with remaining in a curved membrane environment composed of a heterogeneous mixture of lipids. To accommodate the curved topology, the symmetric dimer then undergoes a spontaneous adjustment in which the LH1 ring of one monomer constricts (forming a smaller gap), and this triggers the disorder of its protein-U

(Fig. 7c). Such an event is supported by inspection of the asymmetric distortion of the lipid/micelle belts surrounding the dimeric LH1-RC (Supplementary Fig. 10). Monomer-A has a larger LH1 ring opening than monomer-B and its lipid/micelle belt is discontinued in the gap region (Supplementary Fig. 10b). By contrast, the narrow gap of monomer-B results in a continuous lipid/micelle belt over the junction region (Supplementary Fig. 10d) where four tightly bound CL molecules are present (Fig. 3c). It therefore follows that monomer-B should have the most stable conformation of the two monomers due to more extensive interactions with the surrounding lipids, and as a result, an asymmetric rather than a symmetric dimer forms. In our model, the asymmetric dimer observed in membranes of wild-type *Rba. sphaeroides* thus represents the final mature and functional form of the LH1-RC complex.



**Figure 7 Schematic docking model for the dimerization of LH1-RC complex.** Cartoon representation of the complexes viewed from the periplasmic side. (a) Two C-shaped LH1-RC monomers before dimerization. (b) Dimerization of the monomers at an early stage with a symmetric S-shaped conformation. (c) Formation of an asymmetric junction is completed with breaking symmetry. Due to a wind-up movement of the monomer-B, most parts of protein-U are disordered and the LH1 ring opening is narrowed in the monomer-B. Thus, the sizes of LH1-ring opening are different to enhance asymmetric features. Because monomeric  $\Delta U$  LH1-RC complex lacks protein-U and only has a half LH1 ring, it could not form a stable dimer. The color codes are RC: yellow, LH1 ring: gray, PufX: orange, Protein-U: magenta.

## Methods

**Preparation and Characterization of the Native Dimeric and  $\Delta U$  Monomeric LH1-RC Complexes.** *Rba. sphaeroides* f. sp. *denitrificans* (strain IL106) cells<sup>32</sup> were cultivated phototrophically (anoxic/light) at room temperature for 7 days under incandescent light (60W). Preparation of the native LH1-RC was conducted by solubilizing chromatophores ( $OD_{870\text{-nm}} = 40$ ) with 1.0 % w/v *n*-dodecyl- $\beta$ -D-maltopyranoside (DDM) in 20 mM Tris-HCl (pH 8.0) buffer for 60 min at room temperature, followed by differential centrifugation. The supernatant was loaded onto a DEAE column (Toyopearl 650S, TOSOH) equilibrated at 4 °C with 20 mM Tris-HCl buffer (pH 8.0) containing 0.1 % w/v DDM. The fractions were eluted in an order of LH2, monomeric LH1-RC and dimeric LH1-RC by a linear gradient of NaCl from 0 mM to 400 mM. The dimer-rich LH1-RC fractions were collected and further purified by gel-filtration chromatography (TSKgel G4000SW, 7.5 mm I.D.  $\times$  60 cm, TOSOH) with

20 mM Tris-HCl buffer (pH 7.5) containing 200 mM NaCl and 0.05 % w/v DDM at a flow rate of 1 mL/min at room temperature (Supplementary Fig. 1b). Monomeric LH1-RC lacking protein-U was solubilized from ICM of *Rba. sphaeroides* IL106 strain  $\Delta U^7$  using 1.0 % w/v DDM, followed by sucrose gradient density centrifugation with five-stepwise sucrose concentrations (10, 17.5, 25, 32.5 and 40% w/v) in 20 mM Tris-HCl buffer (pH 8.0) containing 0.05 % w/v DDM at 4 °C and 150,000 $\times$ g for 6 hours. Both native dimeric LH1-RC and monomeric  $\Delta U$  LH1-RC were concentrated for absorption measurement and assessed by negative-stain EM using a JEM-1010 instrument (JEOL).

**Cryo-EM Data Collection.** Proteins for cryo-EM were concentrated to ~5 and 3.4 mg/ml for the native dimer and the  $\Delta U$  mutant, respectively. Three microliters of the protein solution were applied on a glow-discharged holey carbon grids (200 mesh Quantifoil R2/2 molybdenum), which had been treated with H<sub>2</sub> and O<sub>2</sub> mixtures in a Solarus plasma cleaner (Gatan, Pleasanton, USA) for 30 s and then blotted, and plunged into liquid ethane at -182 °C using an EM GP2 plunger (Leica, Microsystems, Vienna, Austria). The applied parameters were a blotting time of 4 s at 80% humidity and 4°C. Data were collected on a Titan Krios (Thermo Fisher Scientific, Hillsboro, USA) electron microscope at 300 kV equipped with a Falcon 3 camera (Thermo Fisher Scientific) (Supplementary Fig. 2). Movies were recorded using EPU software (Thermo Fisher Scientific) at a nominal magnification of 96 k in counting mode and a pixel size of 0.82 Å at the specimen level with a dose rate of 1.03 e- per physical pixel per second, corresponding to 1.25 e- per Å<sup>2</sup> per second at the specimen level. The exposure time was 32.0 s, resulting in an accumulated dose of 40 e- per Å<sup>2</sup>. Each movie includes 40 fractioned frames.

**Image Processing.** All of the stacked frames were subjected to motion correction with MotionCor2<sup>33</sup>. Defocus was estimated using CTFFIND4<sup>34</sup>. All of the picked particles using the crYOLO<sup>35</sup> were further analyzed with RELION 3.1<sup>36</sup>, and selected by 2-D classification (Supplementary Fig. 2, Supplementary Fig. 8, and Supplementary Table 1). Initial 3-D model was generated in RELION, and the particles were divided into four classes by 3-D classification resulting in only one good class. The 3-D auto refinement without any imposed symmetry (C1) produced a map at 2.88 Å and 2.70 Å resolutions for the native dimer and  $\Delta U$  monomer, respectively, after contrast transfer function refinement, Bayesian polishing, masking, and post-processing. Then particle projections were further subjected to subtraction of the detergent micelle density followed by 3-D auto refinement to yield the final map with a resolution of 2.75 Å and 2.64 Å resolutions for the native dimer and  $\Delta U$  monomer, respectively, according to the gold-standard Fourier shell correlation using a criterion of 0.143 (Supplementary Fig. 2, Supplementary Fig. 8 for the native dimer and  $\Delta U$  monomer, respectively)<sup>37</sup>. The local resolution maps were calculated on RESMAP<sup>38</sup>.

**Model Building and Refinement of the Native Dimeric and  $\Delta U$  Monomeric LH1-RC Complexes.** The atomic model of the monomeric LH1-RC complex of *Rba. sphaeroides* (PDB: 7F0L) was fitted to the cryo-EM native dimer and  $\Delta U$  monomer maps, respectively, using Chimera<sup>39</sup>. Real space refinement for the peptides and cofactors were performed using COOT<sup>40</sup>. The manually modified model was refined in real-space on PHENIX<sup>41</sup>, and the COOT/PHENIX refinement was iterated until the refinements converged. Finally, the statistics calculated using MolProbity<sup>42</sup> were checked. Figures were drawn with the Pymol Molecular Graphic System (Schrödinger)<sup>43</sup> and UCSF Chimera<sup>39</sup>.



## REFERENCES

1. Jungas, C., Ranck, J.-L., Rigaud, J.-L., Joliot, P. & Vermeglio, A. Supramolecular organization of the photosynthetic apparatus of *Rhodobacter sphaeroides*. *EMBO J.* **18**, 534-542 (1999).
2. Bahatyrova, S. et al. The native architecture of a photosynthetic membrane. *Nature* **430**, 1058-1062 (2004).
3. Lee, J.K., DeHoff, B.S., Donohue, T.J., Gumpert, R.I. & Kaplan, S. Transcriptional analysis of *puf* operon expression in *Rhodobacter sphaeroides* 2. 4. 1 and an intercistronic transcription terminator mutant *J. Biol. Chem.* **264**, 19354-19365 (1989).
4. Tsukatani, Y. et al. Phylogenetic distribution of unusual triheme to tetraheme cytochrome subunit in the reaction center complex of purple bacteria. *Photosynth. Res.* **79**, 83-91 (2004).
5. Holden-Dye, K., Crouch, L.I. & Jones, M.R. Structure, function and interaction of the PufX protein. *Biochim. Biophys. Acta – Bioenergetics* **1777**, 613-630 (2008).
6. Tani, K. et al. A novel membrane protein in the *Rhodobacter sphaeroides* LH1-RC photocomplex. *bioRxiv*, doi: 10.1101/2021.06.20.449127 (2021).
7. Tani, K. et al. A previously unrecognized membrane protein in the *Rhodobacter sphaeroides* LH1-RC photocomplex. *Nat. Commun.* **12**, 6300 (2021).
8. Qian, P. et al. Cryo-EM structure of the dimeric *Rhodobacter sphaeroides* RC-LH1 core complex at 2.9 Å: the structural basis for dimerisation. *Biochem. J.* **478**, 3923-3937 (2021).
9. Scheuring, S. et al. Structural role of PufX in the dimerization of the photosynthetic core complex of *Rhodobacter sphaeroides*. *J. Biol. Chem.* **279**, 3620-3626 (2004).
10. Siebert, C.A. et al. Molecular architecture of photosynthetic membranes in *Rhodobacter sphaeroides*: the role of PufX. *EMBO J.* **23**, 690-700 (2004).
11. Qian, P., Hunter, C.N. & Bullough, P.A. The 8.5 Å projection structure of the core RC-LH1-PufX dimer of *Rhodobacter sphaeroides*. *J. Mol. Biol.* **349**, 948-960 (2005).
12. Qian, P. et al. Three-dimensional structure of the *Rhodobacter sphaeroides* RC-LH1-PufX complex: Dimerization and quinone channels promoted by PufX. *Biochemistry* **52**, 7575-7585 (2013).
13. Scheuring, S., Busselez, J. & Levy, D. Structure of the dimeric PufX-containing core complex of *Rhodobacter blasticus* by *in situ* atomic force microscopy. *J. Biol. Chem.* **280**, 1426-1431 (2005).
14. Semchonok, D.A., Chauvin, J.-P., Frese, R.N., Jungas, C. & Boekema, E.J. Structure of the dimeric RC-LH1-PufX complex from *Rhodobaca bogoriensis* investigated by electron microscopy. *Phil. Trans. R. Soc. B* **367**, 3412-3419 (2012).
15. Crouch, L.I. & Jones, M.R. 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 (2012).
16. Oelze, J. & Drews, G. Membranes of photosynthetic bacteria. *Biochim. Biophys. Acta* **265**, 209-239 (1972).
17. Wang, Z.-Y., Suzuki, H., Kobayashi, M. & Nozawa, T. Solution structure of the *Rhodobacter sphaeroides* PufX membrane protein: Implications for the quinone exchange and protein-protein interactions. *Biochemistry* **46**, 3635-3642 (2007).
18. Bracun, L. et al. Cryo-EM structure of the photosynthetic RC-LH1-PufX supercomplex at 2.8-Å resolution. *Sci. Adv.* **7**, eabf8864 (2021).
19. Francia, F. et al. 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 (1999).

20. Crouch, L.I., Holden-Dye, K. & Jones, M.R. Dimerization of the *Rhodobacter sphaeroides* RC-LH1 photosynthetic complex is not facilitated by a GxxxG motif in the PufX polypeptide. *Biochim. Biophys. Acta – Bioenergetics* **1797**, 1812-1819 (2010).
21. Ratcliffe, E.C. et al. 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 (2011).
22. Qian, P., Martin, E.C., Ng, I.W. & Hunter, C.N. 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 (2017).
23. Francia, F., Wang, J., Zischka, H., Venturoli, G. & Oesterhelt, D. 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 (2002).
24. Drews, G. & Golecki, J.R. Structure, molecular organization, and biosynthesis of membranes of purple bacteria. in *Anoxygenic Photosynthetic Bacteria* (eds. Blankenship, R.E., Madigan, M.T. & Bauer, C.E.) 231-257 (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995).
25. Hunter, C.N., Pennoyer, J.D., Sturgis, J.N., Farrelly, D. & Niederman, R.A. Oligomerization states and associations of light-harvesting pigment-protein complexes of *Rhodobacter sphaeroides* as analyzed by lithium dodecyl sulfate-polyacrylamide gel electrophoresis. *Biochemistry* **27**, 3459-3467 (1988).
26. Kiley, P.J., Varga, A. & Kaplan, S. Physiological and structural analysis of light-harvesting mutants of *Rhodobacter sphaeroides*. *J. Bacteriol.* **170**, 1103-1115 (1988).
27. Yu, L.-J., Kato, S. & Wang, Z.-Y. Examination of the putative Ca<sup>2+</sup>-binding site in the light-harvesting complex 1 of thermophilic purple sulfur bacterium *Thermochromatium tepidum*. *Photosynth. Res.* **106**, 215-220 (2010).
28. Frese, R.N. et al. Protein shape and crowding drive domain formation and curvature in biological membranes. *Biophys. J.* **94**, 640-647 (2008).
29. Lang, H.P. & Hunter, C.N. The relationship between carotenoid biosynthesis and the assembly of the light-harvesting LH2 complex in *Rhodobacter sphaeroides*. *Biochem. J.* **298**, 197-205 (1994).
30. Gotou, K. et al. Selective expression of light-harvesting complex alters phospholipid compositions in the membranes and core complexes of purple phototrophic bacteria. (*in preparation*).
31. Qian, P. et al. Cryo-EM structure of the monomeric *Rhodobacter sphaeroides* RC-LH1 core complex at 2.5 Å. *Biochem. J.* **478**, 3775-3790 (2021).
32. Satoh, T., Hoshino, Y. & Kitamura, H. *Rhodopseudomonas sphaeroides* forma sp. *denitrificans*, a denitrifying strain as a subspecies of *Rhodopseudomonas sphaeroides*. *Arch. Microbiol.* **108**, 265-269 (1976).
33. Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331-332 (2017).
34. Rohou, A. & Grigorieff, N. Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216-221 (2015).
35. Waner, T. & Raunser, S. The evolution of SPHIRE-crYOLO particle picking and its application in automated cryo-EM processing workflows. *Commun. Biol.* **3**, 61 (2020).
36. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

37. Rosenthal, P.B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721-745 (2003).
38. Kucukelbir, A., Sigworth, F.J. & D., T.H. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63-65 (2014).
39. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-1612 (2004).
40. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of *Coot*. *Acta Cryst.* **D66**, 486-501 (2010).
41. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* **D66**, 213-221 (2010).
42. Chen, V.B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr* **66**, 12-21 (2010).
43. DeLano, W.L. The PyMOL molecular graphics system, DeLano Scientific, LCC, San Carlos, CA USA. (2004).

## ACKNOWLEDGMENT

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. A.T. 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.V.P.N. provided materials, K.T., R. Kanno, R. Kikuchi, S.K., M.H. and A. T. performed the experiments, K.T., R. Kanno, 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-32192 (native dimer) and EMD-32193 ( $\Delta U$  monomer) and PDB-7VY2 (native dimer) and PDB-7VY3 ( $\Delta U$  monomer). All other data are available from the authors upon reasonable request.