Sulfoquinovosyl diacylglycerol is required for dimerization of the *Rhodobacter sphaeroides* RC-LH1 core complex

Elizabeth C. Martin[1], Adam G.M. Bowie[1], Taylor Wellfare Reid[1,3], C. Neil Hunter[1], Andrew Hitchcock[1,*] and David J.K. Swainsbury[2,*]

[1]Plants, Photosynthesis and Soil, School of Bioscience, University of Sheffield, Sheffield, UK.

[2]School of Biological Sciences, University of East Anglia, Norwich, UK.

[3]Present address: Centre for Bacterial Cell Biology, Newcastle University, Newcastle, UK.

*Corresponding authors: Andrew Hitchcock (a.hitchcock@sheffield.ac.uk) and David J.K. Swainsbury (d.swainsbury@uea.ac.uk)

Author ORCID IDs: 0000-0001-9600-7298 (ECM); 0009-0008-6246-4544 (AGMB); 0000-0001-9335-3080 (TWR); 0000-0003-2533-9783 (CNH); 0000-0001-6572-434X (AH); 0000-0002-0754-0363 (DJKS)

Abstract: The reaction centre-light harvesting 1 (RC-LH1) core complex is indispensable for anoxygenic photosynthesis. In the purple bacterium *Rhodobacter* (*Rba.*) *sphaeroides* RC-LH1 is produced both as a monomer in which 14 LH1 subunits form a crescent-shaped antenna around one RC, and as a dimer, where 28 LH1 subunits form an S-shaped antenna surrounding two RCs. The PufX polypeptide augments the five RC and LH subunits, and in addition to providing an interface for dimerization, PufX also prevents LH1 ring closure, introducing a channel for quinone exchange that is essential for photoheterotrophic growth. Structures of *Rba. sphaeroides* RC-LH1 complexes revealed several new components; protein-Y, which helps to form a quinone channel; protein-Z, of unknown function but which is unique to dimers; and a tightly bound sulfoquinovosyl diacylglycerol (SQDG) lipid that interacts with two PufX arginines. This lipid lies at the dimer interface alongside weak density for a second molecule, previously proposed to be an ornithine lipid. In this work we have generated strains of *Rba. sphaeroides* lacking protein-Y, protein-Z, SQDG or ornithine lipids to assess the roles of these previously unknown components in the assembly and activity of RC-LH1. We show that whilst the removal of either protein-Y, protein-Z or ornithine lipids has only subtle effects, SQDG is essential for the formation of RC-LH1 dimers but its absence has no functional effect on the monomeric complex.

Introduction

The purple phototrophic bacterium *Rhodobacter* (*Rba.*) *sphaeroides* contains hundreds of specialised chromatophore vesicles, around 50 nm in diameter, the numbers of which respond to the incident light

intensity [1–3]. The major membrane complexes of the chromatophore comprise the reaction centre-light harvesting 1 (RC-LH1) core complex, the peripheral light harvesting 2 (LH2) antenna complex, the cytochrome (cyt) bc_1 complex, and ATP synthase [4]. Of these, cryogenic electron microscopy (cryo-EM) structures of cyt bc_1 [5], LH2 [6] and both the monomeric and dimeric forms of RC-LH1 [7–10] have recently been determined. Within the chromatophore, LH2 complexes each bind a circular array of bacteriochlorophyll (BChl) and carotenoid pigments that absorb light and transfer the energy via the LH1 complex to the RC, where the energy is transiently stored as a charge separation [11–13]. Light-driven reduction of quinone to quinol at the RC is followed by passage of quinol across the LH1 ring [13] and diffusion to a cyt bc_1 complex, where the quinol is oxidised through operation of the Q-cycle [14–16]. The cyt bc_1 complexes reside in locally lipid-rich domains [4,17] located within a few nm of the RC-LH1 complexes [18]. The catalytic mechanism of the cyt bc_1 complex reduces cytochrome c_2 [11,19], which returns to the photo-oxidised RC, completing the cyclic ET pathway [11,20]; the Q-cycle mechanism also releases protons into the chromatophore lumen, generating a proton-motive force that is utilised by ATP synthase to drive the production of ATP to power cellular metabolism [20,21].

In *Rba. sphaeroides*, the RC is comprised of 3 subunits (L, H and M) and is surrounded in a fixed stoichiometry by a crescent-shaped LH1 ring containing 14 pairs of α and β transmembrane polypeptides, each of which binds two BChls and two carotenoids [7–10,22,23]. Closure of the LH1 ring is prevented by a single copy of the PufX subunit, which precludes the insertion of further LH1 subunits to maintain the open complex [7–10,24–28]. One copy of the recently discovered protein-Y subunit (using the naming convention suggested by Swainsbury *et al* [29]) sits between the RC and the LH1 ring, creating the RC_3-$LH1_{14}$-XY complex and maintaining a separation that allows quinones and quinols to diffuse freely to and from the RC Q_B site.

In purple bacteria most RC-LH1 complexes are monomeric, some with a closed LH1 ring as in *Rhodospirillum rubrum* [30] and *Thermochromatium* (*Tch*.) *tepidum* [31] and others with an incomplete ring - held open by extra scaffolding subunits such as PufX (*Rba. veldkampii*) [32], protein W (*Rhodopseudomonas palustris*) [33–35], or by the insertion of a transmembrane helix associated with the cytochrome subunit (*Roseiflexus castenholzii*) [36,37]. The many structural variations of RC-LH1 complexes are reviewed in [13]. In this context, RC-LH1 complexes in *Rba. sphaeroides* are unusual because they form dimers in which 28 LH1 subunits form an S-shaped assembly around two RCs, which creates a seamless path for energy transfer between the two halves of the complex. This arrangement provides an elegant energy-conservation mechanism that allows an LH1 excited state access to a second RC if the first is already undergoing a charge separation [38,39]. In addition, the dimeric complex is bent with the two monomers held at an angle of 152° [7–10], which imposes curvature on the membranes and is partly responsible for the spherical shape of the chromatophore vesicles [1,8–10,23,40–43]. This property is unique to a few close relatives of *Rba. sphaeroides,* which were recently reclassified into the *Cereibacter* genus [44]. To ensure consistency with the previous literature we will refer to these species as belonging to the *Cereibacter* subgroup and continue to

use the historical species names for *Rba. sphaeroides* and its relatives throughout this manuscript. Mutants of *Rba. sphaeroides* that lack PufX are not only unable to grow photoheterotrophically but no longer form dimers [17,32–35,45,46].

Recent cryo-EM structures of the monomeric and dimeric forms of *Rba. sphaeroides* RC-LH1 have revealed further components in the complex beyond PufX. Protein-Y forms a hydrophobic hairpin structure that lies between the inside surface of LH1α 13 and 14 and the RC. Genetic removal of protein-Y (also called protein U) results in the formation of an incomplete LH1 ring with as few as 11 α and 10 β subunits [9,10]. This suggests that in addition to promoting the access of quinones to the RC Q_B site, protein-Y also provides a binding site for the LH1 subunits at the edge of the LH1 array, ensuring a gap that is correctly positioned to facilitate rapid diffusion of quinone and quinol between the RC and the external quinone pool. Four copies of a second novel transmembrane polypeptide, protein-Z, were also identified in the dimeric structure, but are absent in the monomer, suggesting an as yet undetermined role unique to the dimer [8]. The dimeric structure also revealed the presence of two lipids, the first of which was confidently assigned as sulfoquinovosyl diacylglycerol (SQDG) based on clear density of its distinctly shaped sulfonic acid head group. The other lipid was less well defined but was proposed to be an ornithine lipid [8]. The SQDG lipid was found to bridge the two halves of the dimer, interacting with both the RC L subunit and the critical Arg49 and Arg53 residues of PufX in one monomer, and an LH1 β subunit in the other (Fig.1 A-C). This interaction would explain why mutation of either Arg49 or Arg53 to Lys disrupts dimer formation [47] as this would prevent the interaction between SQDG and PufX that appears to hold the two monomers together.

In this study we investigate the roles of SQDG, ornithine lipids, and the newly identified Y and Z subunits of the RC-LH1 complex by generating a series of mutants deficient in synthesis of specific lipids or lacking the genes that encode protein-Y or protein-Z. The effects of these mutations on RC-LH1 dimerization, RC activity and photoheterotrophic growth have been characterised, providing new insights into these recently identified components of the RC-LH1 core complex.

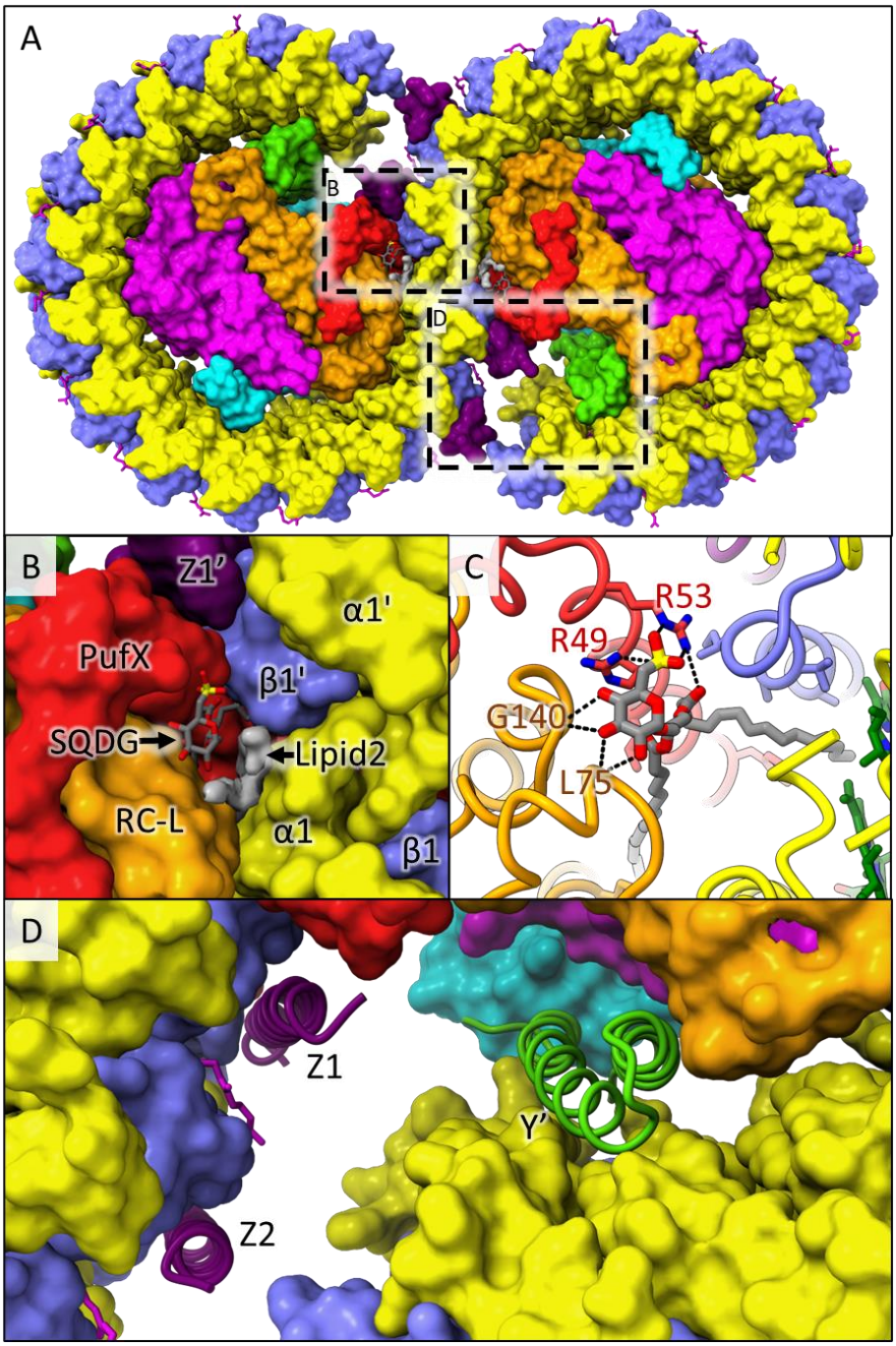


Figure 1. Structure of the dimeric *Rba. sphaeroides* RC-LH1 core complex. (A) Surface view of the complete complex from the lumenal (periplasmic) face. The RC subunits are shown in orange (RC-L), magenta (RC-M) and cyan (RC-H). LH1 subunits are in yellow (α) and blue (β). Additional subunits are in red (PufX), green (protein-Y) and purple (protein-Z). Lipids and cofactors are shown in stick representation in green (BChl), magenta (carotenoids) and grey (SQDG). Unassigned density for lipid 2 is shown as a grey surface. Boxes illustrate areas enlarged in panels B, C and D. (B) Enlarged view highlighting one of the SQDGs and lipid 2 bound at the dimer interface. The subunits from the left monomer (PufX, RC-L, α1 and β1), and those from the right monomer (α1', β1' and Z1') are labelled. (C) A further enlarged view of the SQDG lipid with the protein in ribbon representation. Hydrogen bonds between the SQDG head group and PufX Arg49 and Arg53, and the backbone of RC-L residues Leu75 and Gly140 are shown. (D) Enlarged view of the two protein-Z subunits bound to the left monomer (Z1 and Z2, purple) and protein-Y of the right monomer (Y', green) in ribbon representation. The rest of the protein is in surface representation.

Materials and Methods

Generation of strains and plasmids

The strains used in this study are detailed in Table 1, primer sequences are provided in Table S1, and plasmid information is given in Table S2. Genomic modifications were made using the pK18mob*sacB* plasmid as previously described [48]. Briefly, the target genes were deleted by amplifying ~400 bp regions upstream and downstream of each gene and joining them by overlap extension PCR, yielding a sequence lacking most of the coding region but leaving the start and stop codons intact and in frame to guard against interference with downstream genes upon genomic modification. The resulting PCR fragments were digested with EcoRI and HindIII and ligated into pK18mob*sacB* cut with the same restriction enzymes. The resulting sequence-verified plasmids were transformed into *E. coli* S17-1 and then conjugated into either wild-type or Δ*crtA Rba. sphaeroides.* Correctly modified strains were isolated following sequential selection with kanamycin (30 µg/ml) and counter-selection with sucrose (10% w/v) and the modified loci were verified by PCR and automated Sanger sequencing (Eurofins Genomics).

Expression plasmids were generated by amplifying fragments containing *sqdB* or *sqdBDC* with HindIII and BcuI restriction sites. The digested PCR products were ligated into a modified pBBRBB-P*puf*$_{843\text{-}1200}$ plasmid in which the *puf* promoter-DsRED fragment was replaced with the *pucBAC* genes and 364 bp upstream of *pucB,* corresponding to the *puc* promoter (P*puc*). A HindIII site was engineered immediately downstream of P*puc* such that *pucBAC* could be replaced with a gene of interest by HindIII-BcuI digestion. The *sqdB* and *sqdBDC* plasmids were conjugated into the Δ*sqdB* strain with selection on kanamycin-containing plates, followed by screening of the genomic *sqdB* locus and the gene(s) in the plasmid by PCR.

Table 1. Bacterial strains used in this study.

Species	Description	Source/reference
Escherichia coli		
JM109	Cloning strain for generating plasmid constructs	Promega, UK
S17-1	Conjugative strain for transfer of plasmids to *Rba. sphaeroides*	Simon et al 1983
Rhodobacter sphaeroides		
Wild type	Strain 2.4.1	S. Kaplan*
Δ*crtA*	Unmarked deletion of *crtA* (Rsp_0272); carotenoid pathway truncated at spheroidene	Chi et al 2015 [49]
Δ*puyA*	Unmarked deletion of *puyA* (Rsp_7571); does not produce protein-Y	This study; Qian et al 2021a [7]
Δ*puzA*	Unmarked deletion of *puzA* (orf located within Rsp_2385 gene); does not produce protein-Z	This study; Qian et al 2021b [8]
Δ*olsBA*	Unmarked deletion of *olsBA* (Rsp_3826-3827); does not make ornithine lipids	This study; Aygun-Sunar et al 2006 [50]

Δ*sqdB*	Unmarked deletion of *sqdB* (Rsp_2569); does not make SQDG	This study; Benning and Somerville 1992 [51]
Δ*crtA* Δ*puyA*	Unmarked deletion of *puyA* from Δ*crtA* background	This study
Δ*crtA* Δ*puzA*	Unmarked deletion of *puzA* from Δ*crtA* background	This study
Δ*crtA* Δ*olsBA*	Unmarked deletion of *olsBA* from Δ*crtA* background	This study
Δ*crtA* Δs*qdB*	Unmarked deletion of *sqdB* from Δ*crtA* background	This study
Δ*cycA* Δ*cycI*	Unmarked deletion of *cycA* (Rsp_0296) and *cycI* (Rsp_2577); does not produce cytochrome c_2 (CycA) or isocytochrome c_2 (CycI)	This study

*****Department of Microbiology and Molecular Genetics, The University of Texas Medical School, Houston, TX 77030, U.S.A.

Cell growth and preparation of intracytoplasmic membranes

Rba. sphaeroides cells were grown photoheterotrophically in 1 L Roux bottles containing M22 medium [52] under ~50 µmol m^{-2} s^{-1} illumination from 70 W Phillips Halogen Classic bulbs until they reached stationary phase (an optical density at 680 nm [OD_{680}] of ~3). Strains containing pBBRBB plasmids were supplemented with 30 µg ml^{-1} kanamycin. Cells were harvested by centrifugation at 4,000 RCF for 30 min at 4 °C then resuspended in 5 mL 20 mM Tris-HCl pH 8. Following addition of a few crystals of DNaseI and lysozyme, cells were disrupted via two passes through a chilled French press at 18,000 psi followed by removal of unbroken cells and insoluble debris by centrifugation at 25,000 RCF for 30 min at 4 °C. The supernatant was layered on top of 15/40% (w/v) discontinuous sucrose gradients and centrifuged at 85,000 RCF in a Beckman Type 45 Ti rotor at 4°C for 10 h. A pigmented band of ICM formed at the 15/40% interface, which was harvested using a serological pipette and stored at -20°C until required.

Thin-layer chromatography (TLC) of lipids

TLC was performed based on a method by Swainsbury et al [17] with some modifications. TLC plates were activated by soaking in 0.15 M ammonium sulphate for 15 min and placed in an oven at 160 °C for1 h. 30 µl volumes of membrane fractions, at an optical density at 875 nm (OD_{875}) of 10 were dissolved in 100 µL of 50:50 methanol:chloroform, centrifuged in a benchtop microcentrifuge at 16,000 RCF for 5 min and the lower chloroform phase was removed. 5-10 µl of each lipid standard (~5 mg mL^{-1} in chloroform) and membrane samples were loaded on to the TLC plate and run in 85:15:10:3.5 chloroform:methanol:acetic acid:water for 45 min. The plate was dried for 5 min before being submerged in 50% (v/v) H_2SO_4 for 10-20 s and dried, prior to heating at 160 °C for 10 min to develop the lipid bands.

Fractionation of photosynthetic complexes by rate-zonal centrifugation

Membranes harvested from discontinuous sucrose gradients were diluted at least 5-fold in 20 mM Tris, pH 8 and pelleted at 185,000 RCF for 2 h using a Beckman Type 45 Ti rotor at 4 °C. Pelleted membranes were

resuspended in approximately 100–200 µL of 20 mM Tris pH 8. Six OD_{875} units of resuspended membranes were solubilised in 3% (w/w) n-dodecyl-β-D-maltoside (β-DDM) in a total volume of 375 µl for 1 h at room temperature before centrifugation at 15,000 rpm at 4 °C for 1 h in a microcentrifuge. The supernatant was collected and layered on top of discontinuous sucrose gradients containing steps of 20, 21.25, 22.5, 23.75 and 25% (w/w) sucrose in 20 mM Tris-HCL pH8 and 0.03% (w/v) β-DDM. Gradients were centrifuged in a Beckman SW41 Ti rotor at 125,000 RCF for 40 h at 4 °C. Each gradient was performed in technical triplicate from two biological repeats. Pigmented bands were harvested with a peristaltic pump for downstream processing. If being used for turnover assays, RC-LH1 monomer and dimer bands were buffer exchanged into 50 mM Tris at pH 7.5 with 100 mM NaCl and 0.03% w/v β-DDM by spin concentration with 50,000 MWCO centrifugal concentrators (Sartorius).

To attempt to reform RC-LH1 dimers from monomers of the Δ*sqdB* strain by provision of SQDG, the monomer band harvested from a discontinuous sucrose gradient of the Δ*sqdB* strain was concentrated to 500 µL in a 50,000 MWCO centrifugal concentrator (Sartorius). This sample was split and incubated with and without 0.05% (w/v) SQDG for 24 h before application to another discontinuous sucrose gradient, as described above.

Bioinformatics

Homologues of PufX, protein-Y, protein-Z and the SQDG producing enzyme SqdB were identified by performing tblastn (protein to translated nucleotide) searches against the whole genome contigs (wgs) database on the NCBI BLAST webserver (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Search parameters were calibrated using PufX as a benchmark, which is known to be highly divergent, and were complicated by its short sequence. The settings used to provide positive hits to all PufX proteins in the target species were: Expect threshold = 20; Word size = 2; Max matches in a query range = 0 (default); Matrix BLOSUM62 (default); Gap Costs = Existence: 11, Extension: 1 (default); Compositional adjustments = No Adjustment; Filter Low complexity regions = off. Searches were performed using the *Rba. sphaeroides* PufX sequence as a template (UniProt entry P13402) against each species individually. Top scoring hits for each species were verified by viewing relevant UniProt entries for each sequence, which were annotated as PufX in all cases. Next, searches for LH2 α (UniProt entry Q3J144), protein-Y (UniProt entry U5NME9), the resolved sequence of protein-Z extracted from PDB entry 7PQD, and SqdB (UniProt entry Q3J3A8) were performed using the same settings as for PufX. The top scoring hits from each species were verified by manually inspecting the genomes in the KEGG database (https://www.genome.jp/kegg/) or GENBANK, rejecting sequences embedded within larger genes or in unrelated genomic regions. LH2 hits were also verified via previous experimental determination of LH2 production in these species. A phylogenetic tree of 16S rRNA sequences was generated via a nucleotide BLAST of the *Rba. sphaeroides* 16S RNA (Genbank entry KF791043.1) against the whole genome contigs (wgs) database for all species analysed in this study, and the non-photosynthetic alphaproteobacterium *Caulobacter vibrioides* as an outgroup to root the tree. The tree was rendered using

FigTree v1.44 (available at http://tree.bio.ed.ac.uk/software/figtree/) following flipping of some nodes to sort species by whether they form RC-LH1 dimers. We noted that *puyA* was located near to *otsB* (encoding trehalose-phosphatase) in the genome of the species that have this gene. Whole genome shotgun sequences of *Rba. changlensis* were searched manually for an ORF resembling *puyA* in the vicinity of *otsB.* The sequence in supplementary Figure 6 was found in *Rba. changlensis* strain DSM 18774 NCBI Reference Sequence: NZ_QKZS01000001.1.

Purification of *Rba. sphaeroides* cytochrome c_2

A cytochrome c_2 overproduction strain was constructed by expressing *cycA* under the control of the constitutive P$puf_{842\text{-}1200}$ promoter on the pBBRBB plasmid in a strain lacking the genomic copies of *cycA* and *cycI*. Cell pellets from semi-aerobic cultures (1.6 L of media in 2.5 L conical flasks shaken at 180 rpm at 34 °C) were resuspended in periplasmic extraction buffer (100 mM Tris-HCl pH 8, 500 mM sucrose and 50 mM NaCl) supplemented with a tablet of EDTA-free protease inhibitor (Merck) up to a total volume of 40 mL. 0.8 g of solid sodium deoxycholate (Sigma) was added to the cell resuspension and, after an hour of incubation at 4 °C in the dark, spheroplasts were pelleted at 30,000 RCF for 30 minutes. The supernatant was transferred to a fresh centrifuge tube and 6.25 mL of deoxycholate precipitation solution (1 M Ammonium acetate at pH 5, 250 mM $MgSO_4$) added, before the precipitate was removed by an identical centrifugation step. The supernatant from this second spin step was subsequently passed through 2 0.22µm filters (Sartorius) and made up to 500 ml with 50 mM ammonium acetate buffer at pH 5 before loading onto a 30 mL SP Sepharose column (Cytiva). Cation exchange was performed to purify the cytochrome using a gradient of 13 – 23% buffer B (50 mM ammonium acetate, 1M NaCl).

RC-LH1 turnover assays

Turnover assays were conducted under steady state conditions in a similar fashion to that described in Swainsbury et al., (2021) [23], using 300 µl solutions containing 5 µM reduced cytochrome c_2, 50 µM ubiquinone-2 (Merck) and 0.01-0.05 µM RC-LH1 in a buffer mixture containing 50 mM Tris, 100 mM NaCl, 1 mM sodium-D ascorbate and 0.03 % w/v β-DDM. Following overnight dark adaption, 300 µL of each reaction mixture were placed in a quartz cuvette and monitored at 550 nm using a Cary60 spectrophotometer (Agilent technologies). After 10 seconds, excitation energy was delivered via a fibre optic cable from an 880 nm M810F2 LED (light-emitting diode) (Thorlabs Ltd., UK) driven at 100% intensity using a DC2200 controller (Thorlabs Ltd., UK). The data were processed by fitting the linear initial rate over 0.025-0.1 seconds, starting from the first data point where the absorbance started dropping continuously. Rates were normalised to e^- /RC/sec by dividing the cyt c_2 oxidation rate per second by the RC-LH1 concentration. The concentrations of the RC-LH1 complexes were determined using an extinction coefficient of 3,000 mM^{-1} [53], except for the monomeric Δ*puyA* complex, which we predicted to have an extinction coefficient of 2835 mM^{-1} based upon spectra normalised to the RC bacteriopheophytin band in Fig.3.

Results

Generation and verification of strains lacking SQDG and ornithine lipids

As shown in Fig. 2A, SQDG synthesis from UDP-glucose requires four enzymes: SqdB, SqdA, SqdC and SqdD [51]; we deleted the *sqdB* (Rsp_2569) gene from the *sqdBDC* operon to prevent the first step of SQDG biosynthesis (Fig. 2C). By contrast, just two enzymes, OlsA and OlsB, are required to produce ornithine lipids from L-ornithine [50] (Fig. 2B). The *olsB* (Rsp_3826) and *olsA* (Rsp_3827) genes are arranged into an operon in which the 3' end of *olsB* and the 5' end of *olsA* overlap (Fig. 2D); this gene pair was deleted to prevent ornithine lipid production.

Previous work has shown that the deletion of *crtA* influences the levels of RC-LH1 dimer formation [49]. We therefore deleted *sqdB* and *olsBA* in both a wild-type background and in a strain lacking the spheroidene monooxygenase (Δ*crtA*) to better visualise any effect of lipid content on the monomer-dimer ratio (Fig. 3 A-B). TLC confirmed the loss of SQDG when *sqdB* is deleted (Fig. 2F) but ornithine lipids were not detectable, preventing their analysis using this method. Under photoheterotrophic conditions, growth of the strains unable to produce SQDG or ornithine lipids was indistinguishable from strains with unaltered lipid biosynthesis (see Materials and Methods) (Supplementary Fig. 1), and the UV/Vis/NIR spectra of isolated chromatophore membranes from these strains were also very similar (Supplementary Fig. 2). Therefore, removal of SQDG or ornithine lipids did not result in obvious phenotypes with respect to either growth or spectral features.

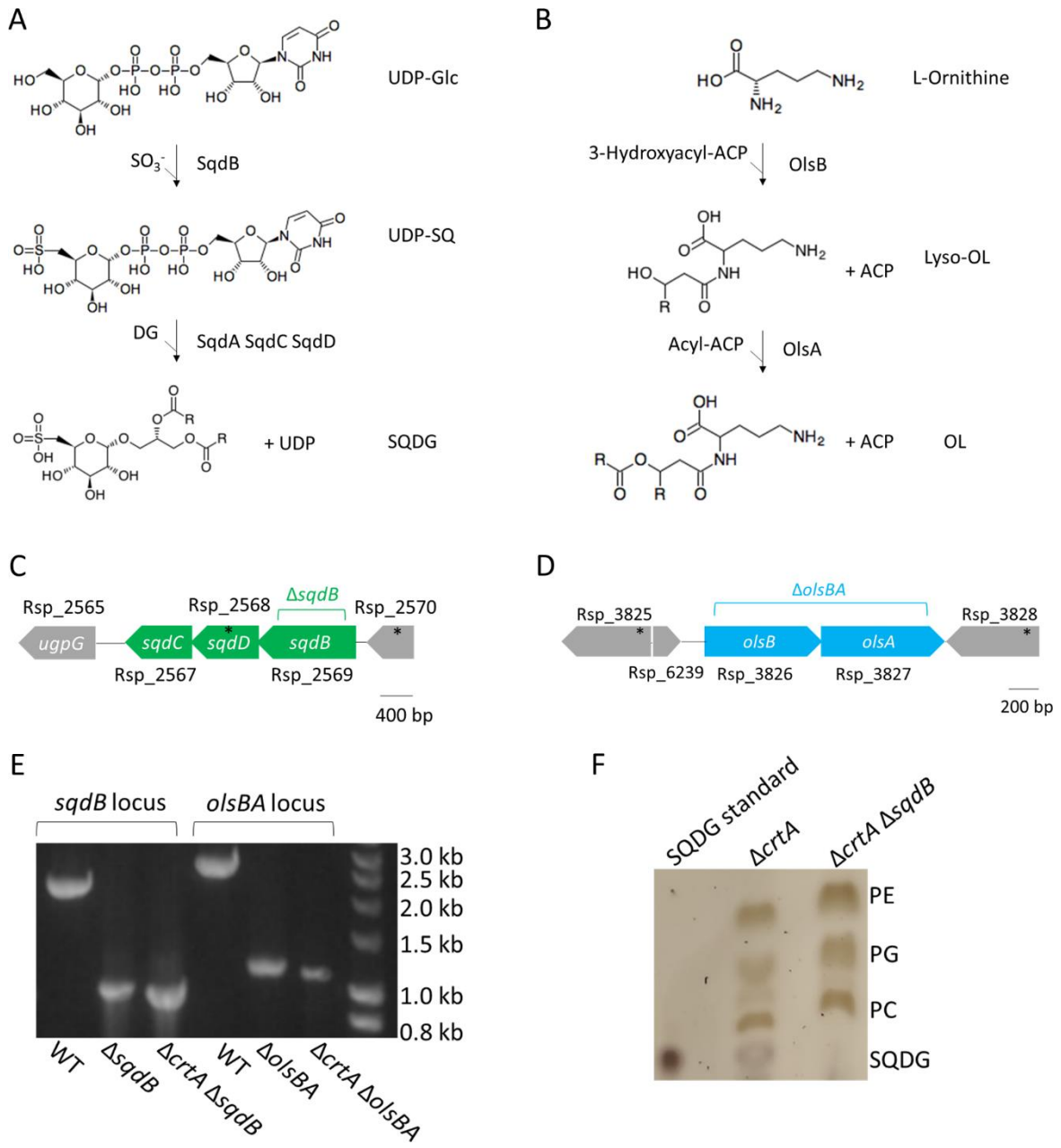


Figure 2. Lipid biosynthesis pathways and deletion of *sqdB* and *olsBA* and PCR verification of knockout strains. (A) sulfoquinovosyl diacylglycerol (SQDG) is synthesised in two steps. The first step is the addition of sulphite to UDP-glucose (UDP-Glc) to produce UDP-sulfoquinovose (UDP-SQ) by SqdB. Next, diacylglycerol is added and UDP is removed by SqdA, C and D to produce SQDG [51]. (B) Ornithine lipids are synthesised by the acyltransferases OlsB and OlsA, which sequentially add a 3-hydroxyacyl group then an acyl group to L-ornithine using 3-hydroxyl-ACP and acyl-ACP as substrates, respectively [54]. (C) Structure of the *sqdBDC* operon. The region labelled Δ*sqdB* (encompassing Rsp_2569) was removed to abolish SQDG biosynthesis. (D) Structure of the *olsBA* operon. The labelled region spanning *olsB* (Rsp_3826) and *olsA* (Rsp_3827) was removed to prevent OL biosynthesis. (E) Agarose gel of ethidium bromide-stained PCR products showing size differences for the amplified regions spanning the *sqdB* and *olsBA* genes showing a clear reduction in size in the knockout strains relative to the wild-type. (F) TLC plate showing loss of SQDG lipid in the Δ*sqdB* strain. A standard for SQDG was run in lane A and a band of the expected size can be seen in samples from Δ*crtA* but not in Δ*crtA* Δ*sqdB* confirming the loss of SQDG biosynthesis.

The loss of SQDG prevents the formation of dimeric RC-LH1 core complexes

The removal of SQDG in wild-type and Δ*crtA* strains results in a complete loss of observable dimer formation, whereas the removal of ornithine lipids has no discernible effect upon the monomer to dimer ratio of RC-LH1 complexes (Fig. 3A-B). The absence of SQDG-free RC-LH1 dimers supports the assignment of SQDG in the cryo-EM structure and highlights the essential role of this lipid for RC-LH1 dimer formation in *Rba. sphaeroides*. The absence of dimers from a control strain known to only produce monomers, in which the PufX Arg49 and Arg53 residues that hydrogen-bond the SQDG headgroup are replaced with leucines [7,8,47], further supports the essential role of SQDG in dimer formation. The normal growth of the cells, and near-identical absorbance spectra of complexes from strains unable to produce SQDG and ornithine lipids (Supplementary Fig. 1 and Fig.3 C-D), suggests that RC-LH1 monomers are properly assembled without these lipids. Introducing *sqdB* to the Δ*sqdB* strain on a replicative plasmid restores some formation of dimers, whilst introducing the complete *sqdBDC* operon fully restores dimer formation to WT levels (Supplementary Fig. 5A). To test whether SQDG could induce dimerization of fully assembled monomeric complexes, we attempted to dimerise monomeric RC-LH1 complexes from the Δ*sqdB* strain by the addition of SQDG. The monomer band harvested from a 20-25% discontinuous gradient of Δ*sqdB* membranes was incubated with or without 0.05% w/v SQDG for 24 hrs, but no significant dimer formation was seen (Supplementary Fig.5B). The inability of exogenously added SQDG to induce dimerization suggests that this lipid has to be fully integrated during the *in vivo* assembly pathway, in order to create the wide range of interactions that stabilise the RC-LH1 dimer. These include a series of hydrogen bonds with the backbone of the RC-L subunit, the salt bridge complex with PufX Arg49 and Arg53, and hydrophobic contacts with the transmembrane region of the opposing LH1β1 subunit and the BChl1 macrocycle on the other side of the complex [8].

To test whether SQDG and ornithine lipids affect RC-LH1 activity *in vitro*, we monitored the light-driven oxidation rate of cytochrome c_2 by monomeric and dimeric complexes in the presence of ubiquinone-2 (an analogue of the native substrate, ubiquinone-10, with a shortened isoprene tail). In monomeric RC-LH1 complexes from the wild-type background, there was no significant difference in turnover rates between the Δ*sqdb* and Δ*olsBA* mutants and equivalent complexes with both lipids, demonstrating no impairment of assembly or function in the absence of either lipid (Fig. 3F). Therefore, we can conclude that whilst deletion of *sqdB* and therefore removal of SQDG biosynthesis precludes dimer formation, there is no functional effect on the monomeric complex.

The absence of proteins Y and Z does not prevent dimer formation.

As the *puyA* ORF (Rsp_7571) encoding protein-Y is isolated in the genome, we were able to excise it without disturbing upstream or downstream genes, leaving behind a sequence encoding 6 residues in the genome of

the unmarked Δ*puyA* strain. The gene encoding protein-Z is located within the Rsp_2385 open reading frame on chromosome 1 (1014511-1014819) but is transcribed in the opposite direction [8]. Most of this gene was deleted to make a Δ*puzA* strain, with just the sequence encoding 7 residues left intact in the genome.

Rate-zonal centrifugation of solubilised chromatophores from these strains show the removal of neither protein-Y nor protein-Z prevents the formation of dimers (Fig.3. A-B). Absorption spectra of each monomer band shows a small but distinct decrease in absorbance at 873 nm in strains lacking protein-Y (Fig 3. C). This observation agrees with the finding that the absence of protein-Y results in monomers missing some α and β LH1 polypeptides, thus they have fewer BChls per RC [9,10]. The dimer bands all have near-identical UV/Vis/NIR spectra, suggesting all dimers contain the same number of α and β polypeptides. The monomer to dimer ratio is similar to wild type in strains lacking either protein-Y or protein-Z, which appear to have no significant role in dimer formation.

Activity assays show that the rate of cytochrome c_2 oxidation by wild-type RC-LH1 monomers and dimers are almost identical, and that rates for both oligomeric forms of the Δ*puyA* RC-LH1 complex are similar, even when accounting for the slight reduction of absorbance at 873 nm (Fig 3E). The activity of the monomeric complexes from the Δ*puzA* strain were found to be similar to the monomeric wild-type complexes, but the activity of the dimeric complexes lacking protein-Z was slightly reduced ($p = 0.03$), suggesting it may have a small, but not essential, functional role exclusive to the dimer (Fig. 3E).

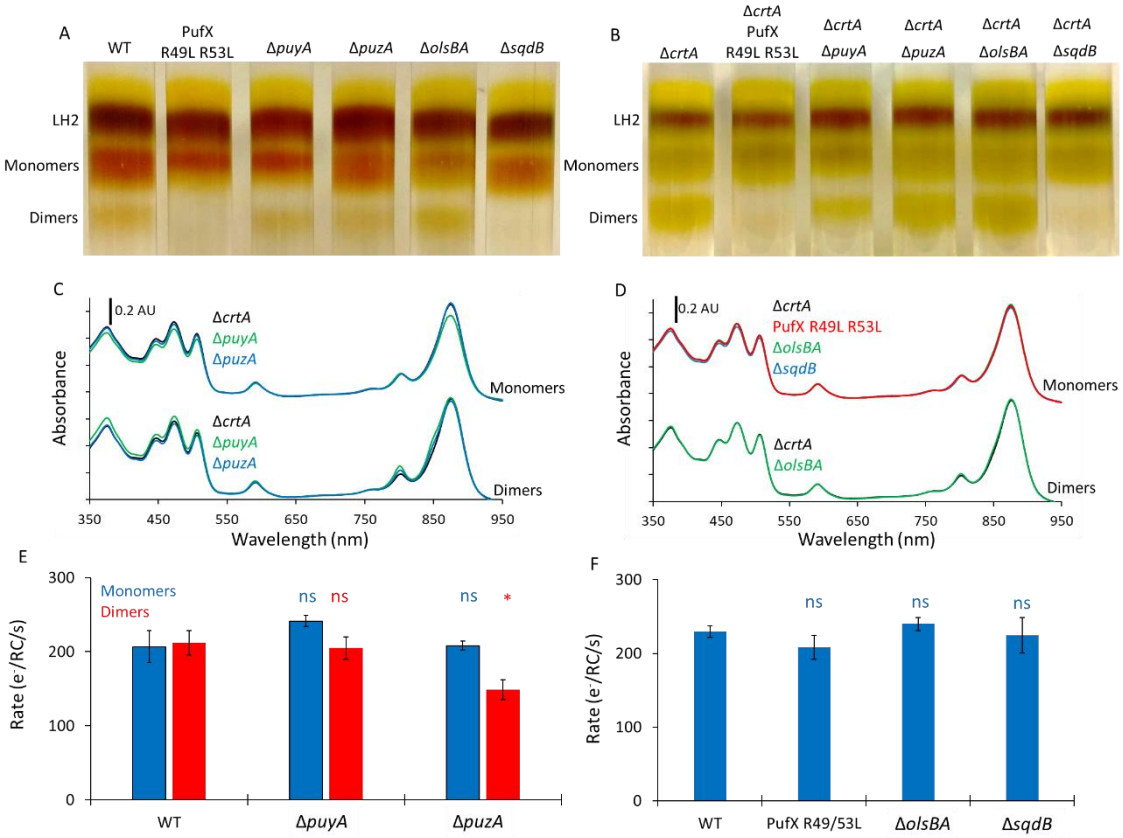


Figure 3. The effects of removing protein-Y, protein-Z, SQDG or ornithine lipids on RC-LH1 dimer formation, absorbance spectra and RC activity. (A-B) Solubilised chromatophore membranes separate into bands of LH2, RC-LH1 monomer and RC-LH1 dimer when centrifuged on sucrose step gradients. Panel A shows the wild-type strain (WT), a control strain that does not produce dimeric RC-LH1 (PufX R49L R53L), a strain lacking protein-Y (Δ*puyA*), a strain lacking protein-Z (Δ*puzA*), a strain that cannot produce ornithine lipids (Δ*olsBA*), and a strain that cannot produce SQDG lipids (Δ*sqdB*). Panel B shows sucrose gradients for the strains in (A) in a background that is also deficient in the *crtA* gene encoding spheroidene monooxygenase (Δ*crtA*). (C-D) UV/Vis/NIR absorbance spectra of the monomer and dimer RC-LH1 bands harvested from the gradients in (B). Panel C shows spectra for the Δ*crtA* strain, and those also harbouring the Δ*puyA* and Δ*puzA* mutations. Panel D shows spectra for the Δ*crtA* strain, and strains also harbouring the RC-LH1 dimer-deficient PufX R49L R53L mutations, and the Δ*olsBA* and Δ*sqdB* genes. (E) Turnover assays for monomeric and dimeric complexes from the WT, and strains lacking proteins Y and Z in strains also lacking CrtA. (F) Turnover assays for the monomeric WT, PufX R49L R53L monomeric control, and monomeric lipid-deficient mutants produced in the wild-type background strain. Rates in E-F show moles of cyt c_2 oxidised per second per mole of RC during illumination using an 810 nm LED of a solution containing 0.01 (E) or 0.05 (F) µM RC-LH1 and 5 µM cyt c_2. T-tests were performed relative to rates for monomeric or dimeric WT complexes where * denotes a p value from 0.01-0.05 and ns denotes a p value > 0.05.

SQDG, PufX, protein-Y and protein-Z in other *Rhodobacter* species

Aligning PufX sequences shows that the two arginine residues that bind SQDG, Arg49 and Arg53, are universally conserved amongst the *Rhodobacter* species within the *Cereibacter* subgroup (Fig.5). All of these species also contain the genes for SQDG biosynthesis (*SqdBDC)*, suggesting RC-LH1 complexes in these species may form dimers with PufX and SQDG in a similar way to *Rba. sphaeroides*. It has been observed that *Rba. azotoformans* and *Rba. changelensis* form dimeric RC-LH1 complexes, but this has yet to be verified for other members of the *Cereibacter* sub-group [55]. The Arg residues are not conserved beyond the *Cereibacter* group and a BLAST search for either *sqdB*, *C* or *D* in the species lacking the PufX Arg residues found no results. We would not expect species with monomeric RC-LH1 complexes to require SQDG, but it is interesting to note that there are two species with dimeric RC-LH1 (*Rhodobaca bogoriensis* [56] and *Rba. blasticus* [57]) that contain neither the Arg49 and Arg53 residues in PufX nor the genes for SQDG biosynthesis. Further exploration is required to see if a different lipid, perhaps with a different mode of binding to PufX, is fulfilling the role of SQDG in dimerization.

We identified sequences for protein-Y in all species within the *Cereibacter* subgroup, and they have extremely high sequence homology that exceeds 94%, except for *Rba. changlensis*, where this value is only 38% (see Fig 4. and Fig. S6 for gene alignments). Protein-Z could also be found in all species in the *Cereibacter* subgroup except for *Rba. changlensis*. The fact that the *Rba. changlensis* PufX sequence is also distinct from the rest of the *Cereibacter* subgroup, and its clear separation in phylogenetic trees generated by its 16S RNA (Fig.4), indicates it may be a more distant relative of *Rba*. *sphaeroides*. We also note that sequence homology was quite high in the first 30 residues of protein-Z, corresponding with the 31 residues resolved in the cryo-EM structure [8], but very low in the rest of the sequence, suggesting function is limited to the N-terminus.

Sequences with homology to *Rba. sphaeroides* protein-Y and protein-Z were not found outside of the *Cereibacter* subgroup. It may be that if proteins are fulfilling the same role, they are not similar enough to be found through searches using sequence homology. This may be the case for the RC-LH1 complex from *Rhodobaca bogoriensis;* structural analysis reveals a protein in a position similar to that adopted by protein-Y in *Rba. sphaeroides* [58], but the DNA sequence identified in the paper bears no homology to *Rba. sphaeroides puyA*.

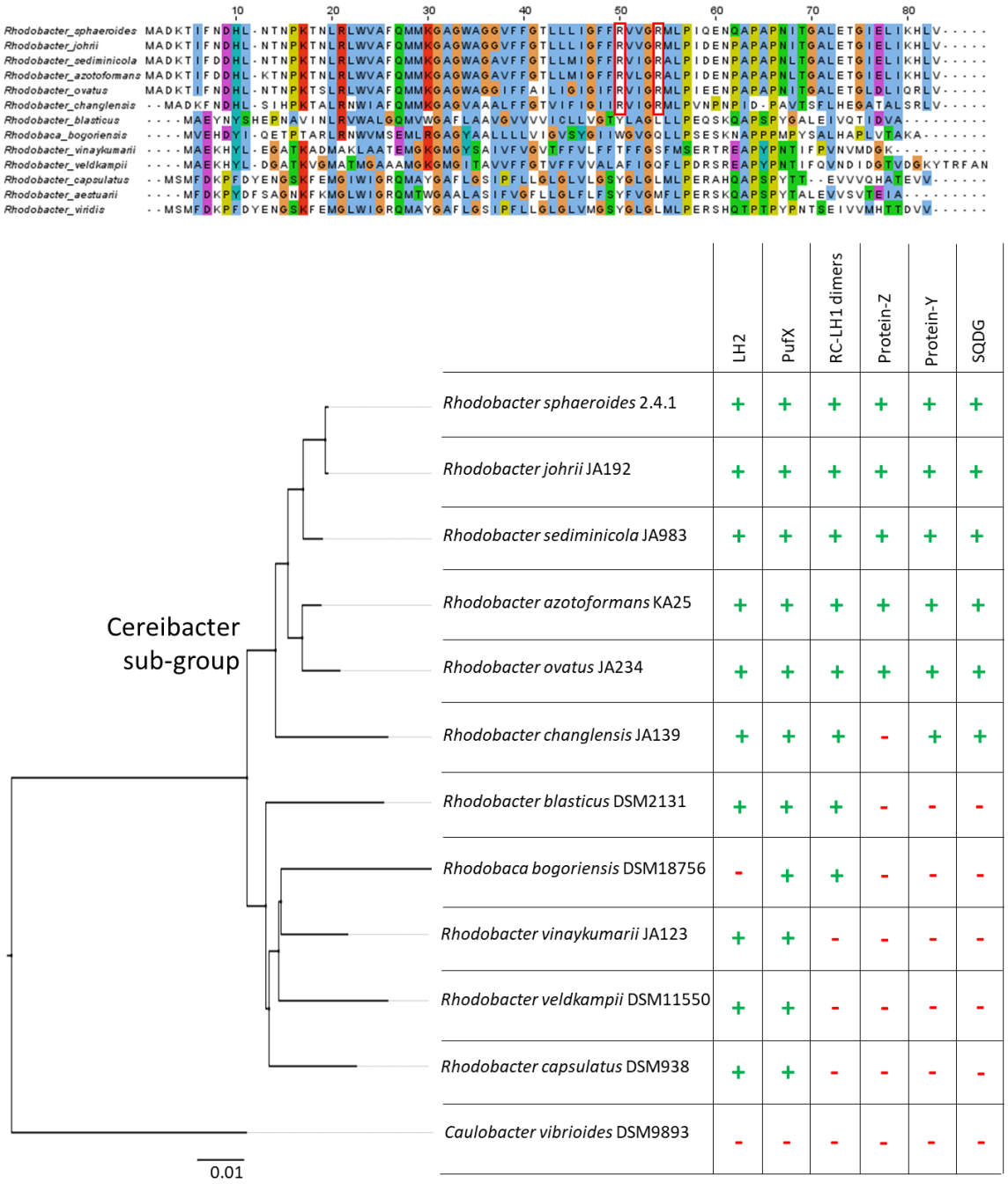


Figure 4. Phylogenetic analysis of the requirements for RC-LH1 dimer formation. (A) Sequence alignment of PufX polypeptides. Red outlined boxes indicate the two arginine residues that bind SQDG in *Rba. sphaeroides*. (B) Left - 16S phylogenetic tree of species that produce PufX for *Rhodobacter* species, and other members of the *Rhodobacteriales* order for which the oligomeric state of their RC-LH1 complex is known. The root of the *Cereibacter* subgroup is labelled. *Caulobacter vibrioides*, a non-photosynthetic alphaproteobacterium, is included as an outgroup to root the tree. Right – A table showing the presence or absence of genes encoding the LH2 complex, PufX, protein-Y and protein-Z, whether the species produces dimeric RC-LH1 complexes, and whether it produces SQDG as inferred by the presence of the *sqdB* gene. Green + symbols indicate presence of a feature and red – symbols indicate its absence.

Discussion

The role of PufX, SQDG and OL in RC-LH1 dimer formation and cellular function

PufX has long been known as a unique component of the RC-LH1 complexes of *Rhodobacter* species [24]. In species that produce dimeric forms of the RC-LH1 complex (recently renamed to *Cereibacter* [44]), PufX is essential for mediating the interaction between pairs of RC-LH1 complexes to form the S-shaped dimers [8–10,23,41,59]. As such, PufX has been the subject of intense investigation to understand how and why certain forms of this polypeptide drive RC-LH1 dimer formation whilst others from closely related species do not [55]. Various elements of PufX have been investigated as potential points of contact between monomers, including a conserved GxxxG motif [27,60], the N-terminal 12 amino acids [26], and a pair of Arg residues at positions 49 and 53 (*Rba. sphaeroides* numbering) that are unique to those PufX polypeptides in dimer forming species [47]. It was not until the determination of high-resolution structures of RC-LH1 dimers that the structure of PufX was modelled with sufficient detail to elucidate the interactions that bring the two RC-LH1 monomers together [8–10]. These structures revealed that the N-terminal 15 amino acids were unresolvable, presumably because they are disordered in the mature complex and only required during dimer assembly. Additionally, a short "LWVAF" motif was observed close to the cytoplasmic surface of the membrane that mediates a direct protein-protein interaction between the two PufX proteins. The GxxxG motif plays no role in bringing together opposing PufX polypeptides. The high-resolution structures also revealed that PufX residues 44-67 were found to bind to the lumenal surface of the reaction centre L subunit.

The large distance between Arg49 and Arg53 in one PufX and those in the opposing monomer was a surprising discovery because these residues are known to be essential for dimer formation. In strains where these Arg residues are replaced with other amino acids [47], including the Δ*pufX* mutant strain we used to determine the structure of the monomeric complex [7], the formation of RC-LH1 dimers is abolished. Careful inspection of the cryo-EM map for the RC-LH1 dimer revealed density consistent with an SQDG lipid for which the head group was hydrogen bonded to R49 and R53. One of the hydrocarbon tails extends between PufX and the L-subunit within the RC-LH1 monomer, and the second tail extends across the dimer interface towards the opposing monomer [8] (Fig.1). The specific nature of the headgroup binding, which could not reasonably accommodate an alternative lipid, and the requirement for R49 and R53 for dimer formation led us to speculate that the SQDG lipid itself is essential for dimer formation. In this study, the observation that removing SQDG from *Rba. sphaeroides* membranes via disruption of the biosynthesis pathway confirms our hypothesis.

In contrast to our findings with SQDG, the removal of OL had no observable effect on dimer formation. This suggests that "lipid 2" (Fig.1 and Qian et al [8]) may not be an OL, or that the lipid bound in this position can be readily substituted for an alternative. Because the density was ambiguous, possibly because of more disorder for lipid 2 relative to SQDG, we are still unable to assign it reliably. We also cannot determine

whether lipid 2 is required for dimer formation. However, its close association with SQDG and its position in a cavity between RC-LH1 monomers would suggest that the presence of a lipid in this location plays an important role.

Surprisingly, removal of SQDG or OL had no observable effect on photosynthetic growth (Fig. S1) or on the activity of the RC-LH1 complexes (Fig. 3E-F). This is consistent with previous observations that disruption of dimer formation does not produce a discernible phenotype under laboratory conditions [47]. This finding also suggests that specific protein-SQDG or protein-OL interactions are not essential in other membrane protein complexes required for growth under photoheterotrophic conditions, although they may be important under other growth modes, as is the case for OL-deficient strains of *Rba. capsulatus* [50]. SQDG is abundant in whole chromatophore membranes and in the annular lipids of complexes extracted using styrene maleic-acid copolymer nanodiscs [17,61,62], so it appears that bound SQDG or OL can often be substituted with other lipids, with the clear exception of the SQDG at the RC-LH1 dimer interface. This is not the first example of essential, ordered lipids being present in RC-LH1 complexes. In most purple bacteria, RCs are known to bind a conserved cardiolipin, the disruption of which via mutation of interacting residues adversely effects RC thermostability [63–65]. There are also many lipids observed between the RC and LH1 ring, many of which are common to all structurally resolved RC-LH1 complexes and interact via conserved residues on the RC and LH1 α subunits [13, 32]. Therefore, this study and those mentioned above highlight the important role of protein-lipid interactions, which are often unresolved in reported RC-LH1 structures.

The role of proteins Y and Z in dimer formation and RC turnover

It had long been assumed that the RC-LH1 complexes of *Rba. sphaeroides* contained five unique proteins (RC-L, RC-M, RC-H, PufX, LH1 α and LH1 β), so the discovery of proteins-Y and -Z in the cryo-EM structures was unexpected [7–10]. Protein-Y was annotated as a hypothetical protein in the UniprotKB, whist protein-Z was unannotated. Our immediate question was whether these newly discovered components of the complex are required for RC-LH1 dimer formation and photosynthetic growth of *Rba. sphaeroides.* Removal of protein-Z had no observable effect on dimer formation or on photoheterotrophic growth but did slightly lower the rates of *in vitro* cytochrome c_2 oxidation by the dimeric complex (p=0.03) (Fig.3B, Fig.S1 and Fig.3E). We note that these assays were only performed in triplicate under a single condition and so are not exhaustive, with further studies required to elucidate the full impact of removing protein-Z. However, these are beyond the scope of the current study.

The removal of protein-Y does not inhibit dimer formation, as indicated by the similar relative abundance of the monomeric and dimeric complexes (Fig. 3A,B). However, the spectra of the monomeric complex show a decrease of LH1 absorbance at 875 nm relative to RC absorbance at 803 and 760 nm, suggesting a lowered LH1 antenna size. Cao *et al* and Tani *et al* recently determined the structure of RC-LH1 lacking protein-Y and found that, whist both monomer and dimer complexes still form, the final subunits of the LH1 antenna fail

to assemble or are dissociated during protein purification [9,10]. Our findings are consistent with the generation of the same complexes with a smaller LH1 antenna. Because protein-Y is distant from the dimer interface and does not interact with PufX or SQDG, it is reasonable that its removal does not affect the ability of the complex to form dimers. Photoheterotrophic growth of the strains unable to produce protein-Y was similar to the wild type, which suggests that the loss of four LH1 subunits has a negligible effect on light-harvesting under our laboratory conditions (Supplementary Fig 1). We previously suspected that protein-Y serves to maintain a channel for efficient quinone exchange, which seems at-odds with our *in vitro* assays. However, loss of the terminal LH1 subunits creates a larger opening in the LH1 ring, which may compensate for the loss of protein-Y at the expense of light-harvesting capacity. Such a loss of capacity would likely not influence growth rates under the illumination conditions used for our experiments.

In strains that lack protein-Z, activity of monomeric complexes was similar to the wild type, which was expected because monomeric RC-LH1 does not bind protein-Z [8]. However, the dimeric complexes that lack protein-Z had slightly lowered activity when compared to wild type dimers (Fig.3E). This suggests that protein-Z may act to ensure the quinone exit channel at the dimer interface is kept open by preventing the complex reverting to the closed state observed by Cao *et al* [9]. Despite the small reduction in activity in purified RC-LH1 dimers lacking protein-Z we could not find an effect on growth rates under our laboratory growth conditions (Supplementary Fig 1).

The evolution of RC-LH1 dimers is synergistic with the presence of SQDG, protein-Y and protein-Z

To further elucidate the roles of SQDG, protein-Y and protein-Z in RC-LH1 dimer formation, we investigated the genomes of other bacteria capable of forming dimeric RC-LH1 complexes to see if they do so via the same mechanism. By searching the relevant databases, we found that the species reclassified into the *Cereibacter* group, all of which are either known to form dimers or we predict will form dimers, appear to contain the genes for SQDG biosynthesis, PufX with SQDG binding residues, protein-Y, and most contain protein-Z, whereas the *Rhodobacter* species that produce monomeric complexes do not appear to contain any of these components. Whilst the phylogenetic analysis we performed is not exhaustive, we are able to speculate that the formation of RC-LH1 dimers in the *Cereibacter* subgroup evolved with the ability to synthesise SQDG, and the evolution of a variant of PufX that could bind the lipid head group. Subsequently, protein-Y was recruited to maintain efficient quinone diffusion. Finally, protein-Z was recruited to lock the complete RC-LH1 dimer in its final conformational state. Exceptions to this are the dimeric RC-LH1 complexes from more distantly related *Rhodobaca bogoriensis* and *Rba. blasticus,* which cannot produce SQDG and appear to be more closely related to the monomer-producing *Rba. capsulatus* and *Rba. veldkampii* than the *Cereibacter* subgroup. At the time of writing the dimeric structure of the *Rhodobaca bogoriensis* complex is published as a preprint [58] and the coordinates are not available in the PDB. However, unlike *Rba. sphaeroides* its LH1 is

473 composed of 15 LH1 subunits and it does not produce LH2, so it may have achieved RC-LH1 dimer formation
474 by a unique mechanism that warrants further investigation.

475

476 **Author contributions:** C.N.H, A.H. and D.J.K.S. conceived and supervised the study. E.C.M., A.G.M.B., T.W.R.,
477 A.H. and D.J.K.S. performed the research. E.C.M., A.G.M.B., C.N.H., A.H. and D.J.K.S. wrote the manuscript.

478

479 **Funding:** A.G.M.B was supported by a University of Sheffield Faculty of Science PhD studentship awarded to
480 C.N.H. C.N.H. acknowledges Synergy Award 854126 from the European Research Council, which partially
481 supported E.C.M. A.H. acknowledges a Royal Society University Research Fellowship (award no.
482 URF\R1\191548), which partially supported E.C.M. D.J.K.S. acknowledges start-up funding from the
483 University of East Anglia.

484

485 **References**

486 1 Tucker, J. D., Siebert, C. A., Escalante, M., Adams, P. G., Olsen, J. D., Otto, C., Stokes, D. L. and
487 Hunter, C. N. (2010) Membrane invagination in *Rhodobacter sphaeroides* is initiated at curved
488 regions of the cytoplasmic membrane, then forms both budded and fully detached spherical
489 vesicles. Mol. Microbiol., Blackwell Publishing Ltd **76**, 833–847.

490 2 Adams, P. G. and Hunter, C. N. (2012) Adaptation of intracytoplasmic membranes to altered light
491 intensity in *Rhodobacter sphaeroides*. Biochim. Biophys. Acta - Bioenerg., Elsevier **1817**, 1616–1627.

492 3 Noble, J. M., Lubieniecki, J., Savitzky, B. H., Plitzko, J., Engelhardt, H., Baumeister, W. and Kourkoutis,
493 L. F. (2018) Connectivity of centermost chromatophores in *Rhodobacter sphaeroides* bacteria. Mol.
494 Microbiol., John Wiley & Sons, Ltd **109**, 812–825.

495 4 Cartron, M. L., Olsen, J. D., Sener, M., Jackson, P. J., Brindley, A. A., Qian, P., Dickman, M. J., Leggett,
496 G. J., Schulten, K. and Hunter, C. N. (2014) Integration of energy and electron transfer processes in
497 the photosynthetic membrane of *Rhodobacter sphaeroides*. Biochim. Biophys. Acta - Bioenerg. **1837**,
498 1769–1780.

499 5 Swainsbury, D. J. K., Hawkings, F. R., Martin, E. C., Musiał, S., Salisbury, J. H., Jackson, P. J., Farmer, D.
500 A., Johnson, M. P., Siebert, C. A., Hitchcock, A., et al. (2023) Cryo-EM structure of the four-subunit
501 *Rhodobacter sphaeroides* cytochrome bc_1 complex in styrene maleic acid nanodiscs. Proc. Natl.
502 Acad. Sci., Proceedings of the National Academy of Sciences **120**, e2217922120.

503 6 Qian, P., Swainsbury, D. J. K., Croll, T. I., Castro-Hartmann, P., Divitini, G., Sader, K. and Hunter, C. N.

(2021) Cryo-EM Structure of the *Rhodobacter sphaeroides* Light-Harvesting 2 Complex at 2.1 Å. Biochemistry, American Chemical Society **60**, 3302–3314.

7 Qian, P., Swainsbury, D. J. K., Croll, T. I., Salisbury, J. H., Martin, E. C., Jackson, P. J., Hitchcock, A., Castro-Hartmann, P., Sader, K. and Hunter, C. N. (2021) Cryo-EM structure of the monomeric *Rhodobacter sphaeroides* RC–LH1 core complex at 2.5 Å. Biochem. J. **478**, 3775–3790.

8 Qian, P., Croll, T. I., Hitchcock, A., Jackson, P. J., Salisbury, J. H., Castro-Hartmann, P., Sader, K., Swainsbury, D. J. K. and Hunter, C. N. (2021) 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.

9 Cao, P., Bracun, L., Yamagata, A., Christianson, B. M., Negami, T., Zou, B., Terada, T., Canniffe, D. P., Shirouzu, M., Li, M., et al. (2022) Structural basis for the assembly and quinone transport mechanisms of the dimeric photosynthetic RC–LH1 supercomplex. Nat. Commun. **13**, 1977.

10 Tani, K., Kanno, R., Kikuchi, R., Kawamura, S., Nagashima, K. V. P., Hall, M., Takahashi, A., Yu, L.-J., Kimura, Y., Madigan, M. T., et al. (2022) Asymmetric structure of the native *Rhodobacter sphaeroides* dimeric LH1–RC complex. Nat. Commun. **13**, 1904.

11 Sener, M., Strumpfer, J., Singharoy, A., Hunter, C. N. and Schulten, K. (2016) Overall energy conversion efficiency of a photosynthetic vesicle. Elife **5**, e09541.

12 Dahlberg, P. D., Ting, P.-C., Massey, S. C., Allodi, M. A., Martin, E. C., Hunter, C. N. and Engel, G. S. (2017) Mapping the ultrafast flow of harvested solar energy in living photosynthetic cells. Nat. Commun. **8**, 988.

13 Swainsbury, D. J. K., Qian, P., Hitchcock, A. and Hunter, C. N. (2023) The structure and assembly of reaction centre-light-harvesting 1 complexes in photosynthetic bacteria. Biosci. Rep. **43**, BSR20220089.

14 Crofts, A. R., Meinhardt, S. W., Jones, K. R. and Snozzi, M. (1983) The role of the quinone pool in the cyclic electron-transfer chain of *Rhodopseudomonas sphaeroides*: A modified Q-cycle mechanism. Biochim. Biophys. Acta **723**, 202–218.

15 Crofts, A. R. (2021) The modified Q-cycle: A look back at its development and forward to a functional model. Biochim. Biophys. Acta - Bioenerg., Elsevier B.V. **1862**.

16 Sarewicz, M., Pintscher, S., Pietras, R., Borek, A., Bujnowicz, Ł., Hanke, G., Cramer, W. A., Finazzi, G. and Osyczka, A. (2021) Catalytic Reactions and Energy Conservation in the Cytochrome bc_1 and b_6f Complexes of Energy-Transducing Membranes. Chem. Rev., American Chemical Society **121**, 2020–

2108.

17 Swainsbury, D. J. K., Proctor, M. S., Hitchcock, A., Cartron, M. L., Qian, P., Martin, E. C., Jackson, P. J., Madsen, J., Armes, S. P. and Hunter, C. N. (2018) Probing the local lipid environment of the *Rhodobacter sphaeroides* cytochrome bc_1 and *Synechocystis* sp. PCC 6803 cytochrome b_6f complexes with styrene maleic acid. Biochim. Biophys. Acta - Bioenerg., Elsevier **1859**, 215–225.

18 Vasilev, C., Swainsbury, D. J. K., Cartron, M. L., Martin, E. C., Kumar, S., Hobbs, J. K., Johnson, M. P., Hitchcock, A. and Hunter, C. N. (2022) FRET measurement of cytochrome bc_1 and reaction centre complex proximity in live *Rhodobacter sphaeroides* cells. Biochim. Biophys. Acta - Bioenerg. **1863**, 148508.

19 Singharoy, A., Barragan, A. M., Thangapandian, S., Tajkhorshid, E. and Schulten, K. (2016) Binding Site Recognition and Docking Dynamics of a Single Electron Transport Protein: Cytochrome c_2. J. Am. Chem. Soc., American Chemical Society **138**, 12077–12089.

20 Singharoy, A., Maffeo, C., Delgado-Magnero, K. H., Swainsbury, D. J. K., Sener, M., Kleinekathöfer, U., Vant, J. W., Nguyen, J., Hitchcock, A., Isralewitz, B., et al. (2019) Atoms to Phenotypes: Molecular Design Principles of Cellular Energy Metabolism. Cell **179**, 1098-1111.e23.

21 Hitchcock, A., Hunter, C. N. and Sener, M. (2017) Determination of Cell Doubling Times from the Return-on-Investment Time of Photosynthetic Vesicles Based on Atomic Detail Structural Models. J. Phys. Chem. B, American Chemical Society **121**, 3787–3797.

22 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.

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 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, American Chemical Society **38**, 6834–6845.

25 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., John Wiley & Sons, Ltd **269**, 1877–1885.

26 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 facilitates dimerisation of the *Rhodobacter sphaeroides* RC–LH1 complex through N-terminal interactions. Biochim. Biophys. Acta - Bioenerg. **1807**, 95–107.

27 Crouch, L. I., Holden-Dye, K. and Jones, M. R. (2010) Dimerisation of the *Rhodobacter sphaeroides* RC–LH1 photosynthetic complex is not facilitated by a GxxxG motif in the PufX polypeptide. Biochim. Biophys. Acta - Bioenerg. **1797**, 1812–1819.

28 Holden-Dye, K., Crouch, L. I. and Jones, M. R. (2008) Structure, function and interactions of the PufX protein. Biochim. Biophys. Acta - Bioenerg. **1777**, 613–630.

29 Swainsbury, D. J. K., Qian, P., Hitchcock, A. and Hunter, C. N. (2023) The structure and assembly of reaction centre-light-harvesting 1 complexes in photosynthetic bacteria. Biosci. Rep. **43**, BSR20220089.

30 Qian, P., Croll, T. I., Swainsbury, D. J. K., Castro-Hartmann, P., Moriarty, N. W., Sader, K. and Hunter, C. N. (2021) Cryo-EM structure of the *Rhodospirillum rubrum* RC–LH1 complex at 2.5 Å. Biochem. J. **478**, 3253–3263.

31 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.

32 Bracun, L., Yamagata, A., Christianson, B. M., Terada, T., Canniffe, D. P., Shirouzu, M. and Liu, L.-N. (2023) Cryo-EM structure of the photosynthetic RC-LH1-PufX supercomplex at 2.8-Å resolution. Sci. Adv., American Association for the Advancement of Science **7**, eabf8864.

33 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., et al. (2021) Structures of *Rhodopseudomonas palustris* RC-LH1 complexes with open or closed quinone channels. Sci. Adv., American Association for the Advancement of Science **7**.

34 Jackson, P. J., Hitchcock, A., Swainsbury, D. J. K., Qian, P., Martin, E. C., Farmer, D. A., Dickman, M. J., Canniffe, D. P. and Hunter, C. N. (2018) Identification of protein W, the elusive sixth subunit of the *Rhodopseudomonas palustris* reaction center-light harvesting 1 core complex. Biochim. Biophys. Acta - Bioenerg. **1859**.

35 Roszak, A. W., Howard, T. D., Southall, J., Gardiner, A. T., Law, C. J., Isaacs, N. W. and Cogdell, R. J. (2003) Crystal Structure of the RC-LH1 Core Complex from *Rhodopseudomonas palustris*. Science (80-.). **302**, 1969–1972.

598 36 Xin, Y., Shi, Y., Niu, T., Wang, Q., Niu, W., Huang, X., Ding, W., Yang, L., Blankenship, R. E., Xu, X., et
599 al. (2018) Cryo-EM structure of the RC-LH core complex from an early branching photosynthetic
600 prokaryote. Nat. Commun. **9**, 1568.

601 37 Qi, C.-H., Wang, G.-L., Wang, F.-F., Xin, Y., Zou, M.-J., Madigan, M. T., Wang-Otomo, Z.-Y., Ma, F. and
602 Yu, L.-J. (2023) New insights on the photocomplex of *Roseiflexus castenholzii* revealed from
603 comparisons of native and carotenoid-depleted complexes. J. Biol. Chem. **299**, 105057.

604 38 Comayras, F., Jungas, C. and Lavergne, J. (2005) Functional consequences of the organization of the
605 photosynthetic apparatus in *Rhodobacter sphaeroides* I. Quinone domains and excitation transfer in
606 chromatophores and reaction center-antenna complexes. J. Biol. Chem. **280**, 11203–11213.

607 39 Timpmann, K., Chenchiliyan, M., Jalviste, E., Timney, J. A., Hunter, C. N. and Freiberg, A. (2014)
608 Efficiency of light harvesting in a photosynthetic bacterium adapted to different levels of light.
609 Biochim. Biophys. Acta - Bioenerg. **1837**, 1835–1846.

610 40 Qian, P., Bullough, P. A. and Hunter, C. N. (2008) Three-dimensional reconstruction of a membrane-
611 bending complex: The RC-LH1-PufX core dimer of *Rhodobacter sphaeroides*. J. Biol. Chem. **283**,
612 14002–14011.

613 41 Siebert, C. A., Qian, P., Fotiadis, D., Engel, A., Hunter, C. N. and Bullough, P. A. (2004) Molecular
614 architecture of photosynthetic membranes in *Rhodobacter sphaeroides*: the role of PufX. EMBO J.
615 **23**, 690–700.

616 42 Frese, R. N., Pàmies, J. C., Olsen, J. D., Bahatyrova, S., van der Weij-de Wit, C. D., Aartsma, T. J., Otto,
617 C., Hunter, C. N., Frenkel, D. and van Grondelle, R. (2008) Protein shape and crowding drive domain
618 formation and curvature in biological membranes. Biophys. J. **94**, 640–647.

619 43 Hsin, J., Gumbart, J., Trabuco, L. G., Villa, E., Qian, P., Hunter, C. N. and Schulten, K. (2009) Protein-
620 induced membrane curvature investigated through molecular dynamics flexible fitting. Biophys. J.
621 **97**, 321–329.

622 44 Hördt, A., López, M. G., Meier-Kolthoff, J. P., Schleuning, M., Weinhold, L.-M., Tindall, B. J., Gronow,
623 S., Kyrpides, N. C., Woyke, T. and Göker, M. (2020) Analysis of 1,000+ Type-Strain Genomes
624 Substantially Improves Taxonomic Classification of Alphaproteobacteria . Front. Microbiol. .

625 45 Tani, K., Kanno, R., Ji, X.-C., Hall, M., Yu, L.-J., Kimura, Y., Madigan, M. T., Mizoguchi, A., Humbel, B.
626 M. and Wang-Otomo, Z.-Y. (2021) Cryo-EM Structure of the Photosynthetic LH1-RC Complex from
627 *Rhodospirillum rubrum*. Biochemistry, American Chemical Society **60**, 2483–2491.

628 46 Walz, T., Jamieson, S. J., Bowers, C. M., Bullough, P. A. and Hunter, C. N. (1998) Projection structures

of three photosynthetic complexes from *Rhodobacter sphaeroides*: LH2 at 6 Å, LH1 and RC-LH1 at 25 Å. J. Mol. Biol. **282**, 833–845.

47 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 - Bioenerg.

48 Sutherland, G. A., Qian, P., Hunter, C. N., Swainsbury, D. J. K. and Hitchcock, A. (2022) Engineering purple bacterial carotenoid biosynthesis to study the roles of carotenoids in light-harvesting complexes. In Methods in Enzymology, Academic Press.

49 Chi, S. C., Mothersole, D. J., Dilbeck, P., Niedzwiedzki, D. M., Zhang, H., Qian, P., Vasilev, C., Grayson, K. J., Jackson, P. J., Martin, E. C., et al. (2015) Assembly of functional photosystem complexes in *Rhodobacter sphaeroides* incorporating carotenoids from the spirilloxanthin pathway. BBA - Bioenerg., Elsevier B.V. **1847**, 189–201.

50 Aygun-Sunar, S., Mandaci, S., Koch, H.-G., Murray, I. V. J., Goldfine, H. and Daldal, F. (2006) Ornithine lipid is required for optimal steady-state amounts of c-type cytochromes in *Rhodobacter capsulatus*. Mol. Microbiol., John Wiley & Sons, Ltd **61**, 418–435.

51 Benning, C. and Somerville, C. R. (1992) Identification of an operon involved in sulfolipid biosynthesis in *Rhodobacter sphaeroides*. J. Bacteriol., American Society for Microbiology **174**, 6479–6487.

52 Hunter, C. N. and Turner, G. (1988) Transfer of genes coding for apoproteins of reaction centre and light-harvesting LH1 complexes to *Rhodobacter sphaeroides*. J. Gen. Microbiol. **134**, 1471–1480.

53 Huang, X., Vasilev, C., Swainsbury, D. J. K. and Hunter, C. N. (2024) Excitation energy transfer in proteoliposomes reconstituted with LH2 and RC-LH1 complexes from *Rhodobacter sphaeroides*. Biosci. Rep. BSR20231302.

54 Vences-Guzmán, M. Á., Geiger, O. and Sohlenkamp, C. (2012) Ornithine lipids and their structural modifications: from A to E and beyond. FEMS Microbiol. Lett. **335**, 1–10.

55 Crouch, L. I. and Jones, M. R. (2012) Cross-species investigation of the functions of the *Rhodobacter* PufX polypeptide and the composition of the RC-LH1 core complex. Biochim. Biophys. Acta - Bioenerg., Elsevier B.V. **1817**, 336–352.

56 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. Philos. Trans. R. Soc. B Biol. Sci., Royal Society **367**, 3412–3419.

57 Scheuring, S., Busselez, J. and Lévy, D. (2005) Structure of the Dimeric PufX-containing Core Complex of Rhodobacter blasticus by in Situ Atomic Force Microscopy. J. Biol. Chem., Elsevier **280**, 1426–1431.

58 Semchonok, D. A., Siponen, M. I., Tüting, C., Charras, Q., Kyrilis, F. L., Hamdi, F., Sadian, Y., Jungas, C. and Kastritis, P. L. (2022) Cryo-EM structure of the *Rhodobaca bogoriensis* RC-LH1-PufX dimeric complex at 2.9. bioRxiv, Cold Spring Harbor Laboratory.

59 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.

60 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.

61 Swainsbury, D. J. K., Scheidelaar, S., Van Grondelle, R., Killian, J. A. and Jones, M. R. (2014) Bacterial reaction centers purified with styrene maleic acid copolymer retain native membrane functional properties and display enhanced stability. Angew. Chemie - Int. Ed. **53**, 11803–11807.

62 Swainsbury, D. J. K., Scheidelaar, S., Foster, N., van Grondelle, R., Antoinette Killian, J. and Jones, M. R. (2017) The effectiveness of styrene—maleic acid (SMA) copolymers for solubilisation of integral membrane proteins from SMA-accessible and SMA-resistant membranes. Biochim. Biophys. Acta - Biomembr. **1859**, 2133–2143.

63 McAuley, K. E., Fyfe, P. K., Ridge, J. P., Isaacs, N. W., Cogdell, R. J. and Jones, M. R. (1999) Structural details of an interaction between cardiolipin and an integral membrane protein. Proc. Natl. Acad. Sci. U. S. A., National Academy of Sciences **96**, 14706–11.

64 Wakeham, M. C., Sessions, R. B., Jones, M. R. and Fyfe, P. K. (2001) Is There a Conserved Interaction between Cardiolipin and the Type II Bacterial Reaction Center? Biophys. J., Elsevier **80**, 1395–1405.

65 Fyfe, P. K., Isaacs, N. W., Cogdell, R. J. and Jones, M. R. (2004) Disruption of a specific molecular interaction with a bound lipid affects the thermal stability of the purple bacterial reaction centre. Biochim. Biophys. Acta - Bioenerg. **1608**, 11–22.

690 **Supporting information**

691 **Table S1. Primers used in this study.** Restriction enzyme sites used for cloning are underlined in bold.

Name	Sequence (5'-3')
olsBA KO Scr F	CTTTCCGAGATCAGCGCCATCTC
olsBA KO UF	GATC**GAATTC**CTGATAAGATCGTGACAGATGCGCG
olsBA KO UR	GACAGGCTCGTCGGCGATCATTCCCGGAC
olsBA KO DF	GATCGCCGACGAGCCTGTCGGCTGACCG
olsBA KO DR	GATC**AAGCTT**GATCGAGAACCATGTGCTGATGGTC
olsBA KO Scr R	GATCGATCTCGAGATCTTCCCCGAC
sqdB KO Scr F	CGGTGGGTGCCGACAAGAT
sqdB KO UF	GATC**GAATTC**GTGGCTGCCATCTGCCAT
sqdB KO UR	CCAATCAGGACACTGCGATGCGCATGAAGCC
sqdB KO DF	CGCATCGCAGTGTCCTGATTGGATCTGGCAG
sqdB KO DR	GATC**AAGCTT**AACAGCCGGTCCACGTTC
sqdB KO Scr R	TCTCGTAGACATTCGGCGCG
puzA KO Scr F	CATTTCTGCATCATCGCGCATGAC
puzA KO UF	CCG**GAATTC**GAAGCTGGACGAGATGTGGAATCC
puzA KO UR	CGTCAGACCTCTTTCATATATGCCATTTAAACCTCCCTCTTGC
puzA KO DF	GAGGTTTAAATGGCATATATGAAAGAGGTCTGACGGACCCGTG
puzA KO DR	CGGC**AAGCTT**CCATCGTTTTCGTCTTCCGTCAC
puzA KO Scr R	GTTCGACGATGGACAGGATCTCG
puyA KO Scr F	CAGCCGATGGTCCAGACCTC
puyA KO UF	CCG**GAATTC**GTCACGATAATGGGCCATGTCTCTC
puyA KO UR	CAGTTGCTGTTTTCGGGCATGGTGCCTCCTTC
puyA KO DF	CCATGCCCGAAAACAGCAACTGACGGCGC
puyA KO DR	CGGC**AAGCTT**GAGGGCTGGATCGACTACGATC
puyA KO Scr R	GGCCTATGTCTCGGGGTTTCTC
cycI KO Scr F	CATTTCGTGAATCCGTCCGAGATCG
cycI KO UF	CCG**GAATTCC**AACGTGAAGGTGATGCGTCAGG
cycI KO UR	CATTTCAGCCCTCCAATCTCATGGTCTTCTCCCTTTGCG
cycI KO DF	GACCATGAGATTGGAGGGCTGAAATGCCTGTCTGC
cycI KO DR	CTG**AAGCTT**GCCCACGTTCTCG
cycI KO Scr R	GCCACAGGATCTTGCCGTCATTG
cycA KO Scr F	CATGGTGGTGAACCTGCAGGAC
cycA KO UF	GCA**GAATTC**CCTCGCATCTGCCGGATACC
cycA KO UR	GGCGACCTGGGCCTTGACTTGGAACTTCATGG
cycA KO DF	GTCAAGGCCCAGGTCGCCGTCCGGC
cycA KO DR	CGC**AAGCTT**GGCGCCTGAATGTACTCACCG
cycA KO Scr R	CTGAAGCAGGCGGTGTCGG
sqdB HindIII F	GATC**AAGCTT**ATGCGCATCGCAGTTCTGG
sqdB BcuI R	GATC**ACTAGT**TCAGGACACCGAGCGCAG
sqdC BcuI R	GATC**ACTAGT**CTAAATCATGAGCGGCAGCGTTTG
sqdB seq F	GTTATCTCGACGTCTCGGTCGAGAC
sqdD seq F	GATTGGATCTGGCAGCCGAAGG
ECM 18	CCTACACGCAAACCGTCGATTTAC

PCR1	CGGGCCTCTTCGCTATT
PCR2	TTAGCTCACTCATTAGG

692

693

694

695 **Table S2. Plasmids used in this study.**

Name	Details	Source/reference
pk18mob*sacB*	Allelic exchange vector, Km^R	Professor J. Armitage*
pk18mob*sacB*::*ΔolsBA*	Construct for unmarked deletion of *olsBA*	This study
pk18mob*sacB*::*ΔsqdB*	Construct for unmarked deletion of *sqdB*	This study
pk18mob*sacB*::*ΔpuzA*	Construct for unmarked deletion of *puzA*	This study
pk18mob*sacB*::*ΔpuyA*	Construct for unmarked deletion of *puyA*	This study
pk18mob*sacB*::*ΔcycA*	Construct for unmarked deletion of *cycA*	This study
pk18mob*sacB*::*ΔcycI*	Construct for unmarked deletion of *cycI*	This study
pBBRBB-P$puf_{843\text{-}1200}$-DsRed	Replicative expression plasmid, Km^R	Addgene.org; Tikh et al 2014
pBBRBB-P$puf_{843\text{-}1200}$-*cycA*	Plasmid for expression of *cycA*	This study
pBBRBB-P*puc-pucBAC*	pBBRBB-P$puf_{843\text{-}1200}$-DsRed with P*puf-Ds*RED replaced with the P*puc-pucBAC*	This study
pBBRBB-P*puc-sqdB*	Plasmid for expression of *sqdB*	This study
pBBRBB-P*puc-sqdBDC*	Plasmid for expression of *sqdBDC*	This study

696 *Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

697

698

699

700

701

702

703

704

705

706

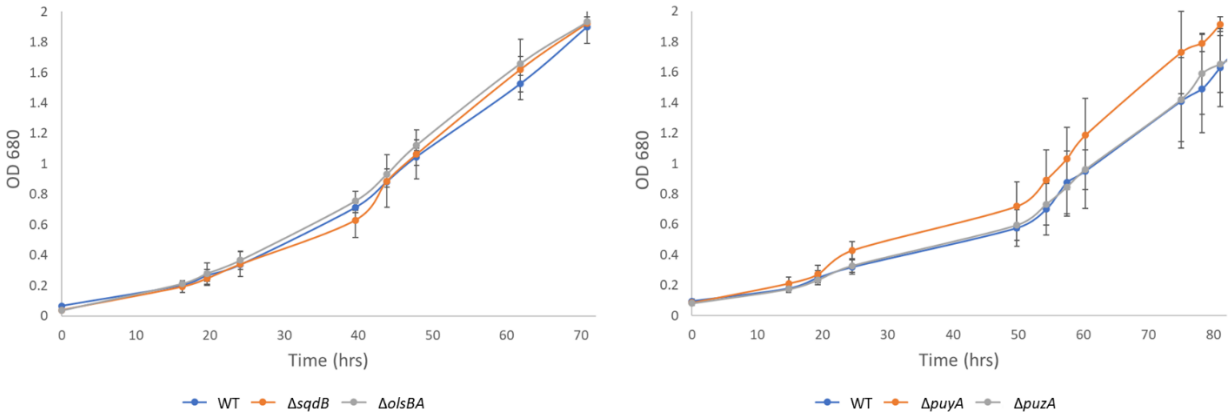
707

708

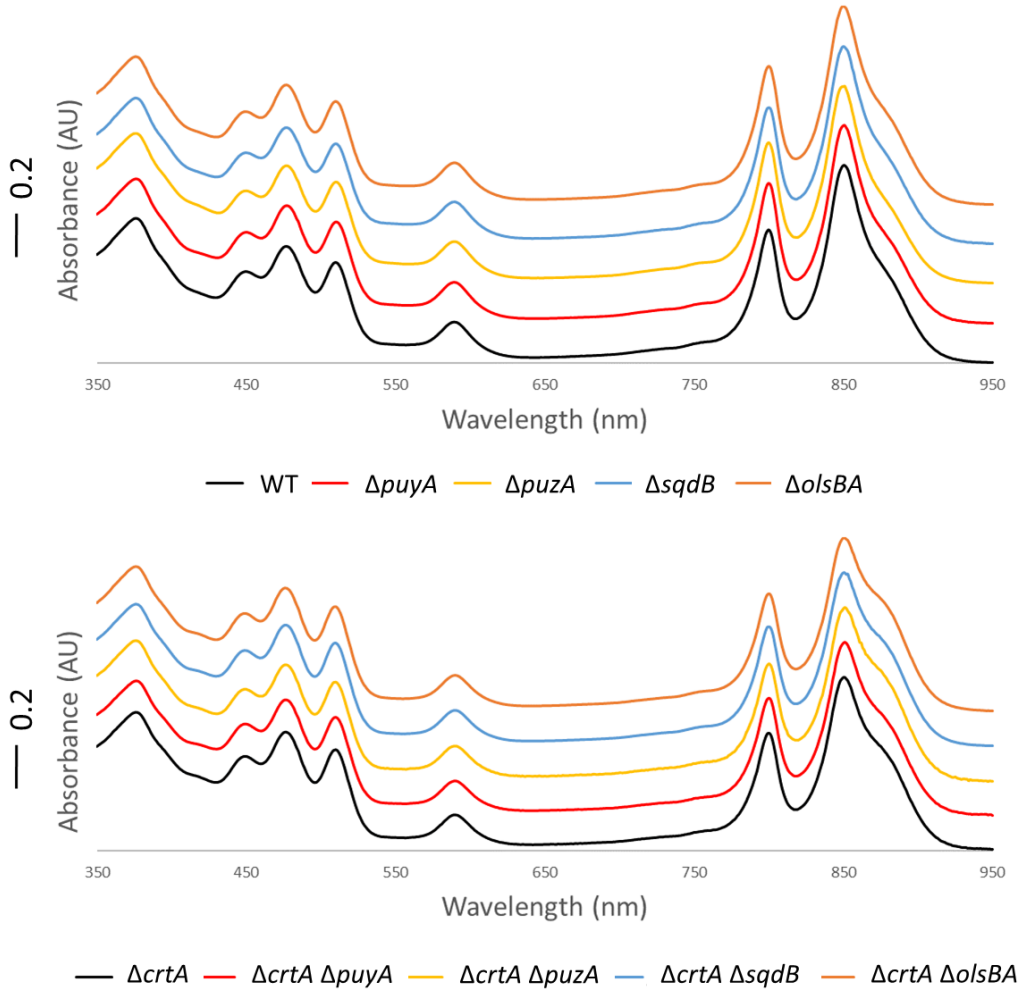
709

710

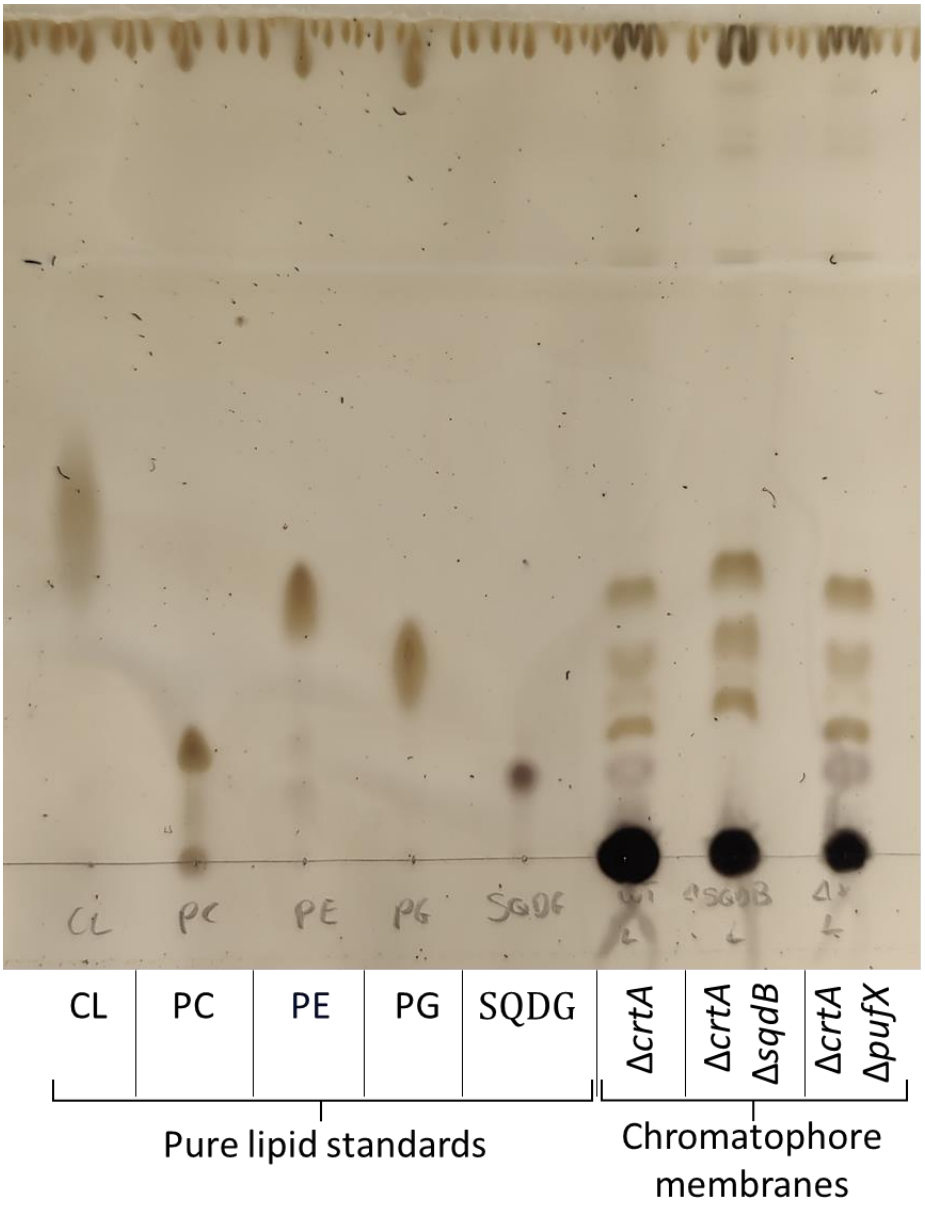
711



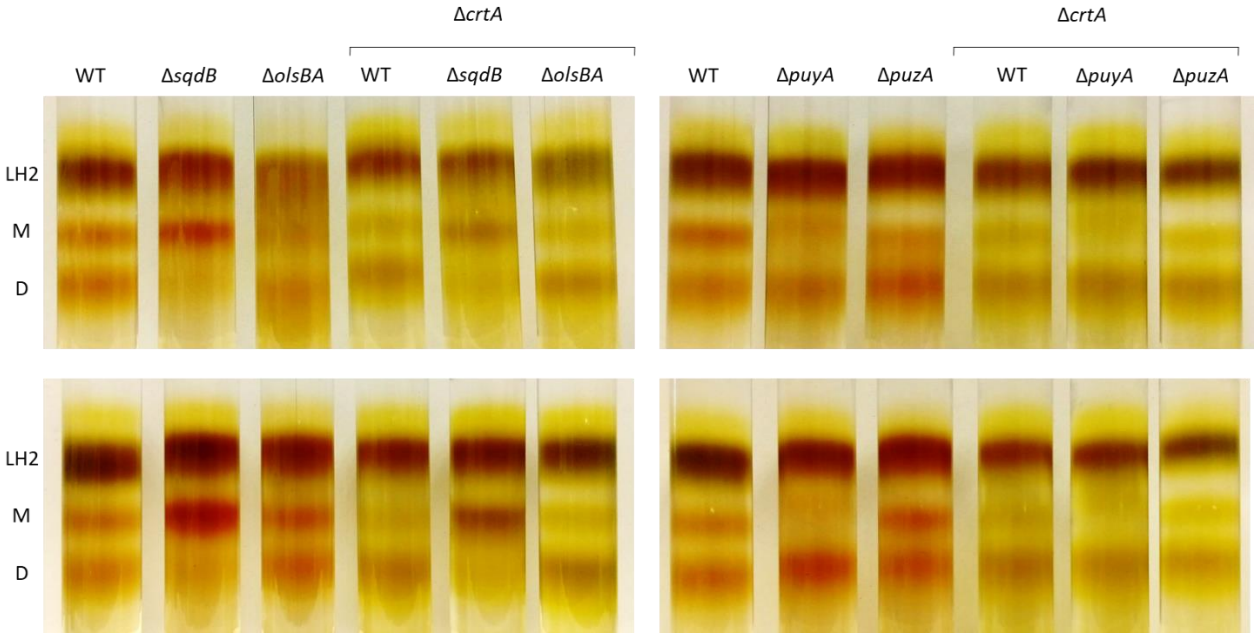
Supplementary Figure 1. Growth curves of all four knockouts in a WT background used in this study at low light (10umol). No significant phenotype was observed in any strain and all knockouts were confirmed by PCR afterwards. Δ*puyA* showed some variation, but further repeats (data not shown) showed the same growth as WT. A phenotype may yet be apparent at different light intensities.



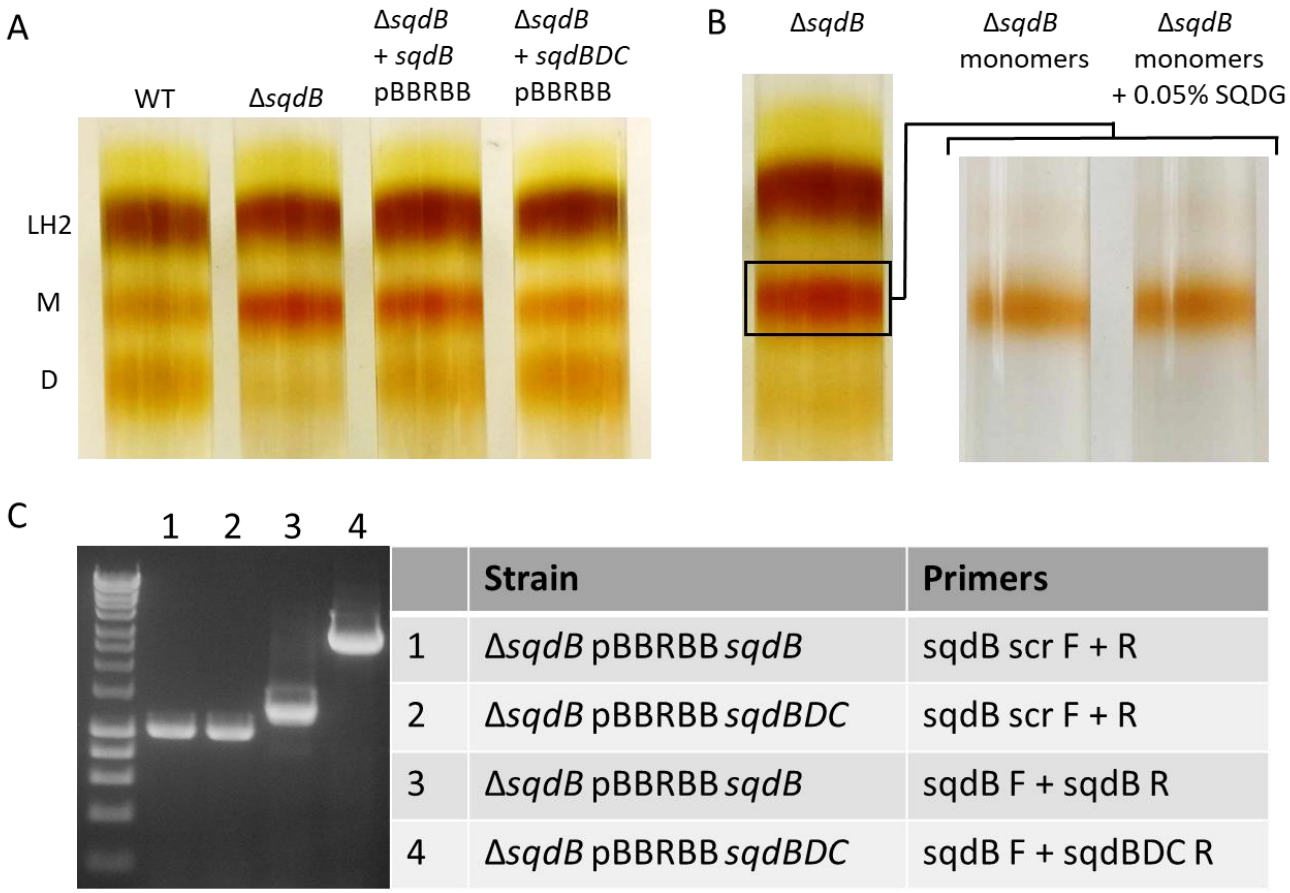
Supplementary Figure 2. UV/Vis/NIR absorbance spectra of chromatophore membranes from all strains in the WT and Δ*CrtA* backgrounds. Spectra collected of chromatophore membranes isolated from other cellular components by differential centrifugation (see methods). Spectra are offset for clarity.



Supplementary Figure 3. Full TLC plate showing pure lipid standards and lipids extracted from chromatophore membranes. The lipid standards were Cardiolipin (CL), Phosphatidylcholine (PC), Phosphatidylethanolamine (PE), Phosphatidylglycerol (PG), Sulfoquinovosyl diacylglycerol (SQDG). Chromatophore membranes were extracted from *Rba. sphaeroides* cells and isolated by separation on 40/15 % w/v sucrose gradients from the Δ*crtA*, Δ*crtA* Δ*sqdB*, and Δ*crtA* Δ*pufX* strains.

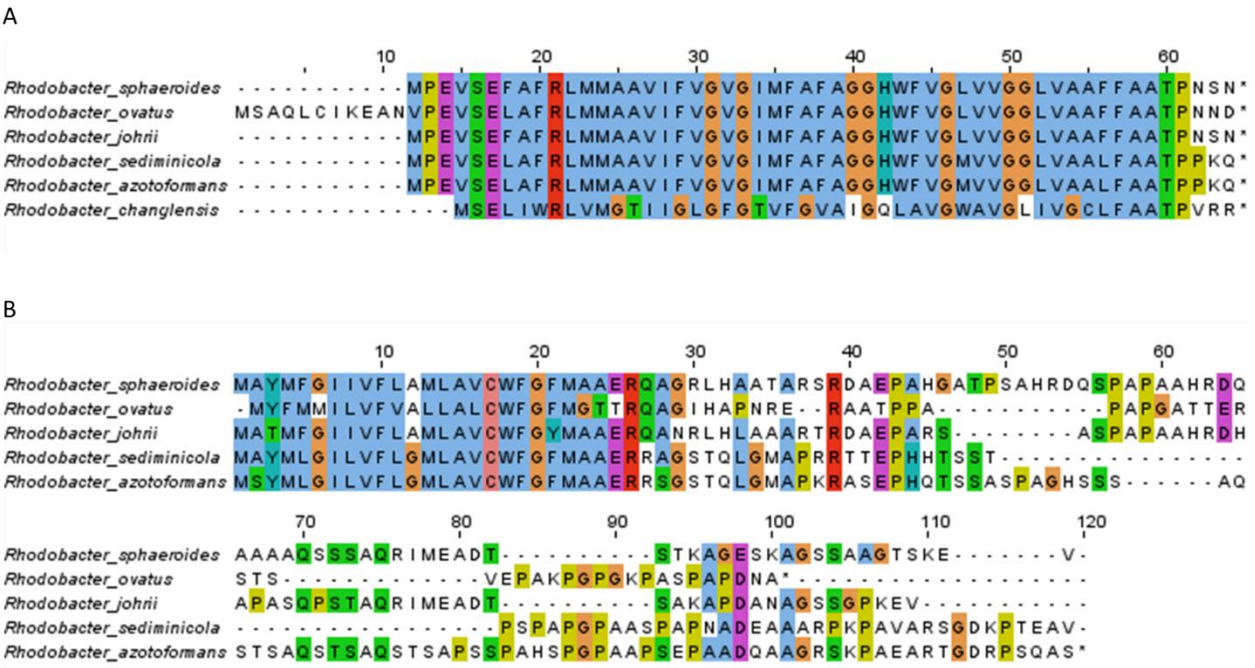


Supplementary Figure 4. Two further repeats were performed showing the same monomer dimer distribution as presented in Figure 3 of the main paper.



	Strain	Primers
1	Δ*sqdB* pBBRBB *sqdB*	sqdB scr F + R
2	Δ*sqdB* pBBRBB *sqdBDC*	sqdB scr F + R
3	Δ*sqdB* pBBRBB *sqdB*	sqdB F + sqdB R
4	Δ*sqdB* pBBRBB *sqdBDC*	sqdB F + sqdBDC R

Supplementary Figure 5. Attempted reconstitution of dimers in the Δ*sqdB* strain by *in trans* complementation or incubation with SQDG. (A) Monomer-dimer gradients of Δ*sqdB* cells expressing *sqdB* from a plasmid show a slight increase in dimer formation. Expression of the *sqdBDC* operon increases dimer expression to WT levels. (B) Purified monomers from Δ*sqdB* incubated with purified SQDG do not spontaneously form dimers. (C) Ethidium bromide-stained PCR products to verify the presence of *sqdB* or *sqdBDC* in pBBRBB-Ppuf$_{843\text{-}1200}$ in the Δ*sqdB* background. Lanes 1 and 2 confirm the absence of *sqdB* in the genome of both strains and lanes 3 and 4 confirm the presence of either *sqdB* or *sqdBDC* on pBBRBB.



Supplementary Figure 6. (A) Alignment of sequences for protein-Y from species within the *Cereibacter* subgroup. (B) Alignments for protein-Z. A sequence for *Rba. changlensis* could not be found, potentially due to a lack of homology. With the exception of *Rba. changlensis*, protein-Y shows a very high degree of sequence homology between species. protein-Z has a disordered tail on the C-terminus that is missing in the structure and shows a very high degree of variation between species. Truncations would have to be performed to establish if this region is unnecessary.