

## PLANT SCIENCES

# Structure of a photosystem I supercomplex from *Galdieria sulphuraria* close to an ancestral red alga

Koji Kato<sup>1†</sup>, Minoru Kumazawa<sup>2†</sup>, Yoshiki Nakajima<sup>1†</sup>, Takehiro Suzuki<sup>3</sup>, Naoshi Dohmae<sup>3</sup>, Jian-Ren Shen<sup>1\*</sup>, Kentaro Ifuku<sup>2\*</sup>, Ryo Nagao<sup>4\*</sup>

Red algae exhibit unique photosynthetic adaptations, characterized by photosystem I (PSI) supercomplexes containing light-harvesting complexes (LHCs), forming PSI-LHCI supercomplexes. In this study, we solved the PSI-LHCI structure of *Galdieria sulphuraria* NIES-3638 at 2.19-angstrom resolution using cryo-electron microscopy, revealing a PSI monomer core associated with seven LHCI subunits. Structural analysis uncovered the absence of phyloquinones, the common secondary electron acceptor in PSI of photosynthetic organisms, suggesting adaptation to a benzoquinone-like molecule. Phylogenetic analysis suggests that *G. sulphuraria* retains traits characteristic of an ancestral red alga, including distinctive LHCI binding and interaction patterns. Variations in LHCI composition and interactions across red algae, particularly in red-lineage chlorophyll *a/b*-binding-like protein and red algal LHCs, highlight evolutionary divergence and specialization. These findings not only deepen our understanding of red algal PSI-LHCI diversification but also enable us to predict features of an ancestral red algal PSI-LHCI supercomplex, providing a framework to explore evolutionary adaptations from an ancestral red alga.

## INTRODUCTION

Oxygenic photosynthesis in cyanobacteria, algae, and land plants harnesses solar energy, converting it into chemical energy with the concomitant production of molecular oxygen (1). This light-driven energy conversion is facilitated by two multisubunit membrane protein complexes, photosystem I (PSI) and PSII, which orchestrate light absorption, charge separation, and electron-transfer processes (2–5). Various light-harvesting antenna subunits bind to the PSI and PSII cores, optimize the capture of solar energy, and facilitate the transfer of excitation energy to the photosystems (1). These antennae exhibit remarkable diversity across photosynthetic organisms in terms of protein sequences and pigment compositions, and are broadly categorized into membrane-bound and water-soluble proteins (1).

The membrane-bound antennae predominantly comprise the light-harvesting complex (LHC) protein superfamily (6, 7), which absorbs solar energy through chlorophylls (Chls) and carotenoids (Cars). Variations in the number and types of Chls and Cars among LHCs contribute to the color diversity observed in photosynthetic organisms, which are broadly classified into green and red lineages (8). The green lineage includes organisms such as green algae and land plants, while the red lineage encompasses red algae, diatoms, haptophytes, cryptophytes, and dinoflagellates (8). LHCs specific to PSI (LHCIs) associate with eukaryotic PSI monomers to form the PSI-LHCI supercomplexes (9, 10), whose structures have been elucidated in various eukaryotic species through cryo-electron microscopy (cryo-EM) (9, 10).

The red-lineage algae have evolved from red algae (11), and the number of LHCLs, as well as their protein and pigment compositions,

exhibits considerable variability (10, 12–14). Red algae constitute a distinct photosynthetic lineage, encompassing both unicellular and large multicellular taxa (15), and serve as an evolutionary intermediate between cyanobacteria and red-lineage eukaryotic algae through a serial endosymbiosis event (8, 11). Recent structural studies have revealed that the red alga *Cyanidioschyzon merolae*, a member of Cyanidiophyceae (Cyanidiophytina) (16–19), features PSI-LHCI structures containing three to five LHCI subunits (20, 21). Similarly, the red alga *Cyanidium caldarium*, a member of Cyanidiophyceae (16–19), has a PSI-LHCI structure containing five LHCI subunits, as observed in the strain NIES-2137 (22). In contrast, the red alga *Porphyridium purpureum*, belonging to Porphyridiophyceae, an order included in Rhodophytina, which is a group different from Cyanidiophyceae, displays a PSI-LHCI structure with seven LHCI subunits and a red-lineage Chl *a/b*-binding-like protein (RedCAP) (23), which is part of the LHC protein superfamily (6, 7). These observations underscore notable differences in both the number and binding sites of LHCLs across these three red algal species.

The red alga *Galdieria sulphuraria* belongs to Galdieriales, an order that likely retains more ancestral red algal traits than other Cyanidiophyceae, such as Cavernicolales, Cyanidiales, and Cyanidioschyzonales (19). Phylogenetic analysis robustly supported that Galdieriales diverged earliest, followed by Cavernicolales, which diverged before Cyanidiales and Cyanidioschyzonales (19). Furthermore, Cyanidiales and Cyanidioschyzonales have undergone substantial genome reduction (24). The shared characteristics of Galdieriales with Rhodophytina—the other group of red algae outside of Cyanidiophyceae—can provide key information on the trajectory of early red algal evolution. In particular, the structural arrangement of LHC around PSI in Galdieriales is predicted to provide insights into the assemblage of the ancestral red algal structure (25). While negative-stain electron microscopy has suggested features of the PSI-LHCI supercomplexes in *G. sulphuraria* (26), the overall structure has yet to be determined.

In this study, we solved a PSI-LHCI structure of *G. sulphuraria* NIES-3638 at a resolution of 2.19 Å by cryo-EM single-particle analysis. The structure shows a PSI monomer core and seven LHCI subunits.

Copyright © 2025 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

<sup>1</sup>Research Institute for Interdisciplinary Science and Graduate School of Environmental, Life, Natural Science and Technology, Okayama University, Okayama 700-8530, Japan. <sup>2</sup>Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan. <sup>3</sup>Biomolecular Characterization Unit, RIKEN Center for Sustainable Resource Science, Saitama 351-0198, Japan. <sup>4</sup>Faculty of Agriculture, Shizuoka University, Shizuoka 422-8529, Japan.

\*Corresponding author. Email: shen@cc.okayama-u.ac.jp (J.-R.S.); ifuku.kentaro.2m@kyoto-u.ac.jp (K.I.); nagry@shizuoka.ac.jp (R.N.)

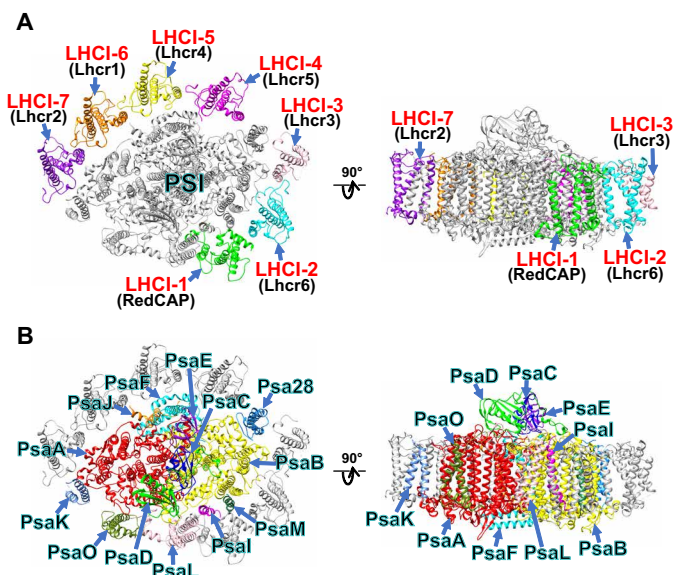
†These authors contributed equally to this work.

On the basis of the structural and phylogenetic analyses of LHCI, we propose an ancestral red algal PSI-LHCI structure and provide evolutionary insights into the conservation and diversity of red-lineage LHCI.

## RESULTS

### Overall structure of the *G. sulphuraria* PSI-LHCI supercomplex

The PSI-LHCI supercomplexes were purified from the red alga *G. sulphuraria* NIES-3638, and their biochemical and spectroscopic properties have been characterized previously (27). Cryo-EM images were obtained using a JEOL CRYO ARM 300 electron microscope operated at 300 kV, yielding a final cryo-EM map at 2.19-Å resolution with C1 symmetry (figs. S1 and S2 and table S1). The atomic model of PSI-LHCI was built on the basis of this cryo-EM map (see Materials and Methods, fig. S2, and tables S1 to S3). The structure reveals a monomeric PSI core bound with seven LHCI subunits (Fig. 1, A and B). The seven LHCI subunits are designated LHCI-1 to LHCI-7 (Fig. 1A) following the naming convention established for LHCI subunits in the PSI-LHCI structure of *C. caldarium* RK-1 (NIES-2137) (22); among which, the names and binding sites of LHCI-1 and LHCI-2 are conserved between the *G. sulphuraria* and *C. caldarium* PSI-LHCI structures. The PSI core of *G. sulphuraria* includes 99 Chls *a*, 22  $\beta$ -carotenes (BCRs), 5 zeaxanthins (ZXTs), 3 [4Fe-4S] clusters, 2 unknown quinone molecules, and 6 lipid molecules, whereas the 7 LHCI subunits include 69 Chls *a*, 25 ZXTs, 4  $\beta$ -cryptoxanthins (BCXs), 2 BCRs, and 5 lipid molecules (table S3).



**Fig. 1. Overall structure of the PSI-LHCI supercomplex from *G. sulphuraria*.** Structures are viewed from the stromal side (left) and the direction perpendicular to the membrane normal (right). Only protein structures are shown, and cofactors are omitted for clarity. The LHCI (A) and PSI core (B) subunits are labeled and colored differently. The seven LHCI subunits are labeled as LHCI-1 to LHCI-7 (red), with their gene products indicated in parentheses (black) in (A). "Lhcr" stands for the light-harvesting complex subfamily originally found in red algae, and "RedCAP" for the red lineage chlorophyll *a/b*-binding-like protein.

### Structure of the *G. sulphuraria* PSI core

The PSI core comprises 13 subunits: PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaI, PsaJ, PsaK, PsaL, PsaM, PsaO, and Psa28 (Fig. 1B). The arrangement of these subunits, excluding Psa28, is conserved among the three species of *G. sulphuraria*, *C. merolae*, and *C. caldarium* in Cyanidiophyceae (21, 22). Notably, all 13 subunits in the *G. sulphuraria* PSI-LHCI (GsPSI-LHCI) structure closely match those in the PSI-LHCI structure of *P. purpureum* in Porphyridiophyceae of Rhodophytina (23). Encoded by the nuclear gene *Psa28*, the Psa28 subunit features two transmembrane helices that interact with PsaB. Psa28 was initially identified in the diatom *Chaetoceros gracilis* (28, 29) and has also been observed in the PSI core of various red-lineage algae, including the rhodophyte *P. purpureum* (23), cryptophytes *Chroomonas placoides* (30), *Rhodomonas salina* (31), dinoflagellates *Symbiodinium* sp. (32, 33), and *Amphidinium carterae* (33). Psa28 is also referred to as PsaR; however, the name of Psa28 was adopted here on the basis of the historical background of nomenclature of PSI proteins, as explained in our recent study (34).

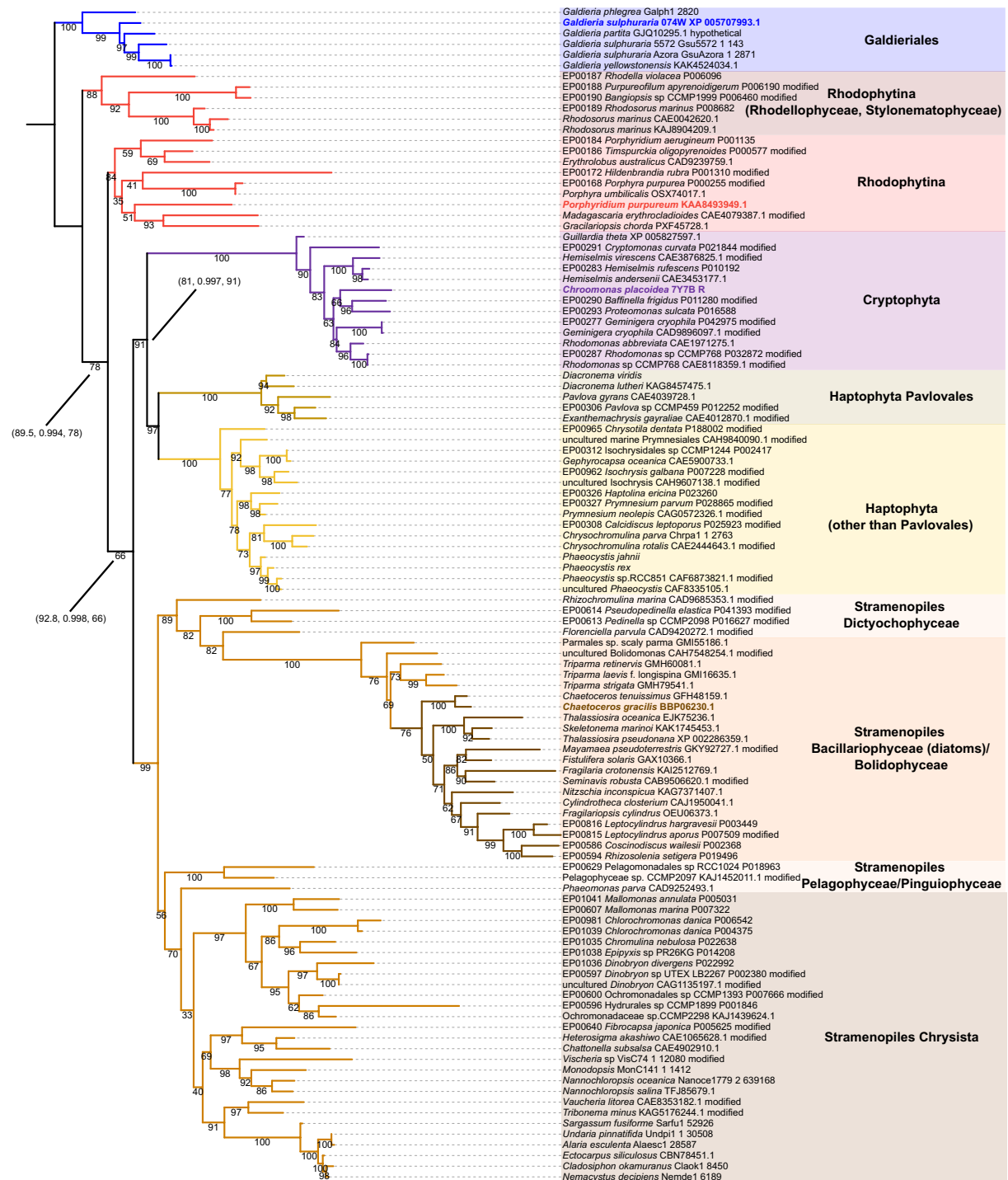
### Molecular evolution and significance of Psa28

Despite the widespread presence of Psa28 among red-lineage algae, a comprehensive phylogenetic analysis of Psa28 is yet to be conducted. Here, we address this gap by collecting Psa28 sequences from a broad range of red-lineage algae. Psa28 is conserved in the Galdieriales of Cyanidiophyceae and in Rhodophytina, which includes *P. purpureum*. However, Psa28 is absent in Cyanidiales (including *C. caldarium*) and Cyanidioschyzonales (including *C. merolae*) within Cyanidiophyceae, while the presence of Psa28 in Cavernicolales, an order of Cyanidiophyceae, remains uncertain because of the lack of reported nuclear genome data (19).

In the secondary endosymbiotic algae of red lineage, Psa28 is present in cryptophytes, stramenopiles (including diatoms), haptophytes, and dinoflagellates. To identify conserved motifs, we aligned these Psa28 sequences, which reveal that while the two transmembrane helices of Psa28 are highly conserved, two regions display taxon-specific variations, namely, (i) the N-terminal region interacting with PsaB (fig. S3, A and B) and (ii) the structurally diverse luminal region (fig. S3, C to E). In the N-terminal region, T52, W54, and G56 of *G. sulphuraria* Psa28 interact with E303, A307, and R309 of PsaB at distances of 2.6 to 3.0 Å, respectively (fig. S3A). These residues of Psa28 are conserved not only in other Galdieriales species but also in some Rhodophytina species, e.g., *Gracilariopsis chorda*, *Porphyra umbilicalis*, *Porphyra purpurea*, *Madagascaria erythrocladioides*, and *Rhodella maculata* (fig. S3B), but not in *P. purpureum*, which would use other residues for PsaB-Psa28 interactions. However, considering the conservation between *G. sulphuraria* and some Rhodophytina species, the PsaB-Psa28 interaction in *P. purpureum* may be acquired secondarily. Therefore, it is reasonable to assume that Psa28 is bound to PSI via PsaB, as it was already present in their last common ancestor and has been vertically inherited to Galdieriales and Rhodophytina. Notably, diatoms and Bolidophyceae (synonymously, Parmales) within stramenopiles have a unique protein motif in the luminal loop region of Psa28 (a green square in fig. S3C), whereas cryptophytes lack both of this motif and the additional residues corresponding to PLTV present in *G. sulphuraria* (a red square in fig. S3C).

We performed a molecular phylogenetic analysis using Psa28 sequences from a broad range of lineages (Fig. 2), excluding dinoflagellate sequences due to low similarity. The Psa28 clades of the secondary endosymbiotic algae of red lineage are nested within the

Tree scale: 1



**Fig. 2. Molecular phylogenetic tree of Psalmodium 28S rDNA.** The phylogenetic tree was inferred by IQ-TREE 2 and rooted at the out-group of the Galdieriales RedCAP clade based on the species tree of red algae (19, 71). Numbers on nodes are ultrafast bootstrap support (%). Numbers in parentheses are SH-like approximate likelihood ratio test (SH-aLRT) support (%), aBayes support, and ultrafast bootstrap support (%) values, respectively.



Rhodophytina clade, forming a monophyletic group. In this large clade, haptophyte and cryptophyte Psa28 sequences also form a monophyletic group, with an ultrafast bootstrap value of 91%. Similarly, Psa28 sequences of heterokonts (photosynthetic stramenopiles) show strong monophyly, with a 99% bootstrap value. Unlike in the RedCAP phylogeny (25), however, heterokonts and haptophytes do not form a monophyletic group in the Psa28 phylogeny.

In the Psa28 phylogeny, the secondary endosymbiotic algae of red lineage branch from the Rhodophytina clade, contrasting with the plastid phylogeny, where red-lineage secondary endosymbiotic algae and Rhodophytina form sister groups (35). Nevertheless, the Psa28 phylogenetic tree resembles the plastid genome tree, where each secondary endosymbiotic alga of red lineage, such as cryptophytes and haptophytes, forms monophyletic groups. This topology of the Psa28 phylogenetic tree and the conserved PsaB–Psa28 interaction suggest that Psa28 should be inherited vertically from the common ancestor of red algae to Galdieriales and Rhodophytina and horizontally to the red-lineage secondary endosymbiotic algae from red algae. Despite Psa28 being nuclear-encoded, this resemblance to plastid phylogeny may reflect evolutionary constraints driven by its role as a PSI core protein. Specifically, the strong interactions between Psa28 and the adjacent plastid-encoded protein PsaB (fig. S3A) may act as a selective force, maintaining this phylogenetic pattern. These findings offer insights into the evolution and functional adaptation of Psa28 across diverse red-lineage algae.

### Unique quinone molecules of the *G. sulphuraria* PSI core

The electron-transfer chain of the *G. sulphuraria* PSI comprises the special pair Chl P700, accessory Chls A<sub>CC</sub>, primary electron acceptors A<sub>0</sub>, secondary electron acceptors A<sub>1</sub>, and three iron-sulfur clusters F<sub>X</sub>/F<sub>A</sub>/F<sub>B</sub> (Fig. 3A). Most of the electron-transfer chain components of *G. sulphuraria* are identical to those of *C. caldarium* (22); however, the secondary electron acceptors could not be identified as phyloquinones, which are present in PSI of most photosynthetic organisms, including cyanobacteria, algae, and land plants. This discrepancy arises from F<sub>0</sub> – F<sub>c</sub> difference maps (red meshes) of the two A<sub>1</sub> molecules analyzed by Servalcat, showing negative difference densities at the ring structures between phyloquinone models (cyan sticks) and observed maps (blue meshes) (Fig. 3B). The details of how the quinone models fit the maps are summarized in fig. S4. Ubiquinone-4 was well fitted into the maps corresponding to the two A<sub>1</sub> molecules, with both the two methoxy groups (black arrows) and a methyl group (red arrows) within their ring structures being clearly resolved (fig. S4, A and B). However, phyloquinone did not fit into the maps at the positions indicated by black arrows (fig. S4, C and D), with apparent steric clashes between the models and the maps. Plastoquinone-9, which is bound to the quinone sites in PSII, lacked a corresponding structural feature in the models at the positions indicated by red arrows (fig. S4, E and F), preventing a satisfactory fit into the maps. High-performance liquid chromatography (HPLC) analysis showed no apparent phyloquinone peak in the PSI core of *G. sulphuraria* (black line in Fig. 3C), with reference to the HPLC chromatogram of pure phyloquinone (red line in Fig. 3C). Similarly, ubiquinone-4 was undetected in the PSI core compared with the reference (green line in Fig. 3C). Instead, a peak for an unidentified quinone appeared at a retention time of 32.9 min (black line in Fig. 3C). The absorption spectrum of this unknown quinone differed from those of phyloquinone and ubiquinone-4 (Fig. 3D) but was more similar to that of the ubiquinone-4 than that of the

phyloquinone. The other detectable peaks corresponded to pigments derived from Chls or Cars (fig. S5). On the basis of these findings, we provisionally modeled the two A<sub>1</sub> molecules as ubiquinone-4, which fit the cryo-EM map, including its ring and side chain (Fig. 3, A and B, and fig. S4).

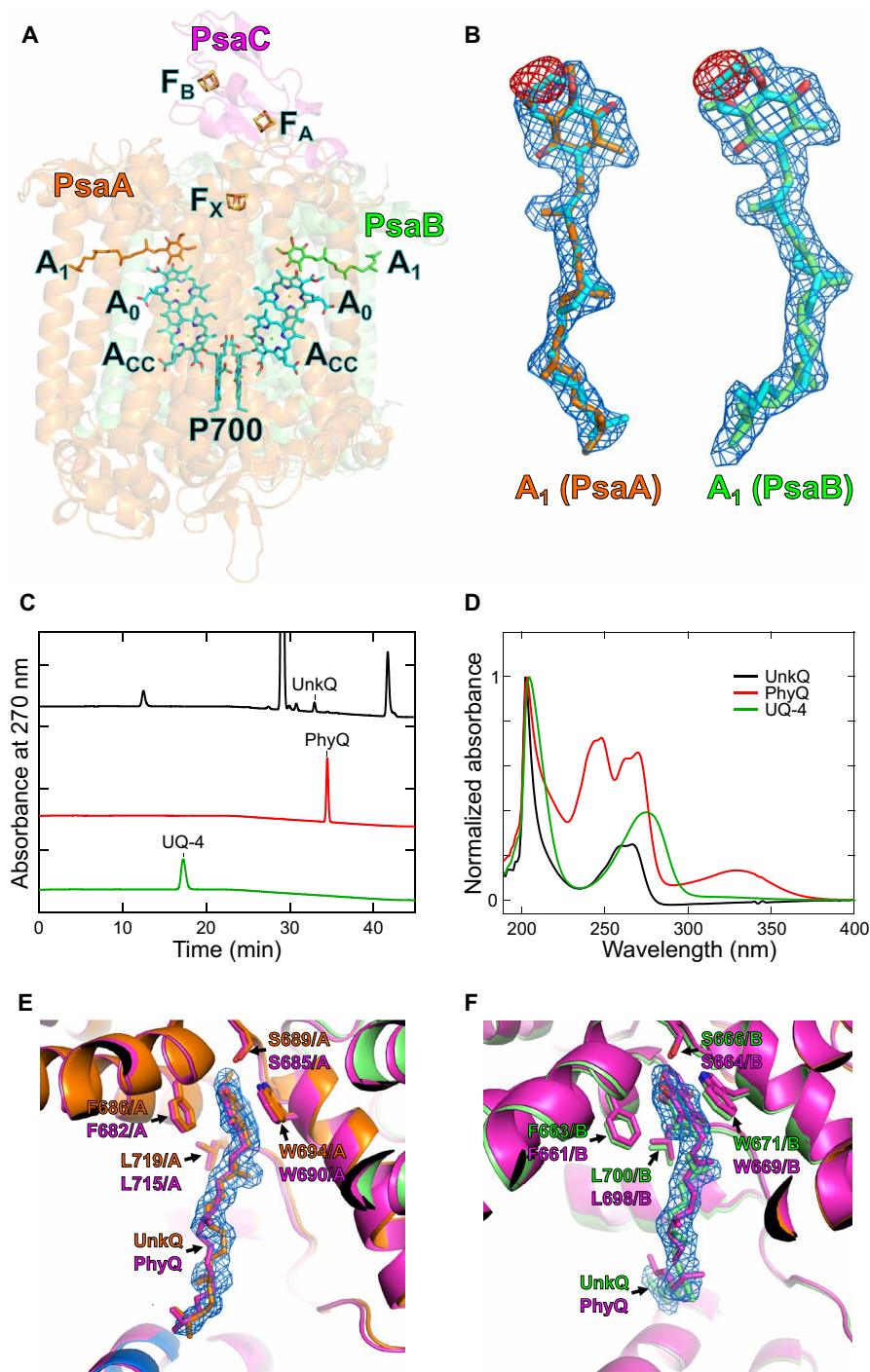
Both oxygenic and anoxygenic photosynthetic organisms typically use naphthoquinones such as phyloquinone or structurally similar menaquinone-4 as A<sub>1</sub> molecules in PSI and PSI-type reaction centers (36). Green plants use phyloquinone as A<sub>1</sub> in PSI, as do most cyanobacteria and the red alga *P. purpureum* (37, 38). By contrast, the cyanobacterium *Gloeobacter violaceus*, the red alga *C. caldarium* IAM R-11 (NIES-2137), and the diatom *C. gracilis* have been reported to use menaquinone-4 as A<sub>1</sub> (39–41). Phyloquinone and menaquinone-4, though differing in their side chains, share a naphthoquinone core and exhibit similar electrochemical properties. In cryo-EM structural analyses, distinguishing phyloquinone from menaquinone-4 is highly challenging, and the presence of phyloquinone cannot be excluded in the PSI cores of *C. caldarium* NIES-2137 and *C. gracilis*; consequently, phyloquinone has been modeled in the structures of *C. caldarium* NIES-2137 PSI-LHCI and *C. gracilis* PSI-FCPI (22, 28, 29). Ubiquinone, with its benzoquinone structure commonly found in respiratory chains, has different electrochemical properties, including a higher redox potential than naphthoquinones (42), making it less ideal as an A<sub>1</sub> electron acceptor. In addition, ubiquinone-4 was not found in the PSI structures from various photosynthetic organisms (9, 10, 43). Therefore, it is interesting how the benzoquinone-like compound observed in the *G. sulphuraria* PSI may exhibit similar redox potential to phyloquinone and menaquinone-4.

Notably, the surrounding environment of A<sub>1</sub>, including protein structures and amino acid residues of PsaA and PsaB in the PSI-LHCI structure of *G. sulphuraria*, is nearly identical to that of *C. caldarium* NIES-2137 (Fig. 3, E and F, and figs. S6 and S7). The amino acid sequences of PsaA and PsaB had similarities of 94 and 92%, respectively, between *G. sulphuraria* and *C. caldarium* NIES-2137 (figs. S6 and S7). The side chain length of the unknown quinone modeled as ubiquinone-4 matches that of the phyloquinone, and they differ only in their ring structures (Fig. 3B). This suggests that phyloquinone- and ubiquinone-4-like molecules can bind to the A<sub>1</sub> sites irrespective of their differences in ring structures, as no distinctive cavities exist in the protein environment surrounding the A<sub>1</sub> ring structures (Fig. 3, E and F).

The presence of different quinones at the A<sub>1</sub> sites in *G. sulphuraria* may provide insights into the origin of quinones in the red algal PSI cores. The cluster of the menaquinone biosynthesis genes in the plastid genome in Cyanidiophyceae is suggested to be acquired through horizontal gene transfer from Chlamydiae, with the exception of Galdieriales (19). However, on the basis of current data, the origin and biosynthesis of the benzoquinone-like compound in GsPSI-LHCI remain uncertain. Further investigation into the chemical identity and biosynthetic pathways of this benzoquinone-like compound is necessary to better understand the diversity and origins of A<sub>1</sub> in red algae and other photosynthetic organisms.

### Pigment composition of the *G. sulphuraria* PSI core

The number and arrangement of Chls within the PSI core in the GsPSI-LHCI structure (fig. S8A) closely resemble those in the *C. caldarium* PSI-LHCI structure (22). However, Chl a102 of PsaI and Chl a201 of Psa28 are unique to the GsPSI-LHCI structure. In contrast, the Car composition of PSI in *G. sulphuraria* (fig. S8B) differs



**Fig. 3. Characteristics of the A<sub>1</sub> molecules.** (A) Arrangement of cofactors involved in the electron-transfer reaction of GsPSI-LHCI. P700, special pair Chl; A<sub>CC</sub>, accessory Chl; A<sub>0</sub>, primary electron acceptor; A<sub>1</sub>, secondary electron acceptor; F<sub>X</sub>, F<sub>A</sub>, and F<sub>B</sub>, iron-sulfur clusters. (B) Cryo-EM maps and quinone structures (sticks) of the A<sub>1</sub> molecules. The A<sub>1</sub> molecules of GsPSI-LHCI were provisionally modeled as ubiquinone-4 in PsaA (orange) and PsaB (green). Two ubiquinone-4 molecules superimposed with phyloquinones (cyan). Red meshes, F<sub>0</sub> - F<sub>C</sub> difference maps ( $-10\sigma$ ) when the A<sub>1</sub> molecules are modeled as phyloquinones; blue meshes, denoised maps ( $3\sigma$ ). (C) HPLC profiles of the *G. sulphuraria* PSI core (black line), phyloquinone (PhyQ; red line), and ubiquinone-4 (UQ-4; green line). The *G. sulphuraria* PSI cores were prepared previously (27) and showed a peak of unknown quinone (UnkQ). PhyQ and UQ-4 were purchased (see Materials and Methods). (D) Absorption spectra of UnkQ (black line), PhyQ (red line), and UQ-4 (green line). The spectra were measured by a photodiode array detector in the HPLC system and normalized by the maximum peak intensity of each spectrum. Data in [(C) and (D)] are representative of three independent experiments. (E and F) Structures of PsaA (E) and PsaB (F) that surround the A<sub>1</sub> molecules. The GsPSI-LHCI structure (orange for PsaA and green for PsaB) was superimposed with the PSI-LHCI structure of *C. caldarium* NIES-2137 [Protein Data Bank (PDB): 8WEY, magenta for both PsaA and PsaB]. Amino acid residues participating in the interactions with the ring of A<sub>1</sub> are labeled; for example, F686/A and F663/B mean Phe<sup>686</sup> of PsaA and Phe<sup>663</sup> of PsaB, respectively.

from that in *C. caldarium* (22), with ZXT853 of PsaB, ZXT206/ZXT207 of PsaO, and BCR202/ZXT203 of Psa28 found only in *G. sulphuraria*. BCR304 of PsaF in *G. sulphuraria* is replaced by ZXT205 of PsaF in *C. caldarium*.

### Structure of the *G. sulphuraria* LHCI

The seven LHCI subunits in the PSI-LHCI structure were identified as the products of seven genes: RedCAP, Lhcr6, Lhcr3, Lhcr5, Lhcr4, Lhcr1, and Lhcr2, at the sites of LHCI-1 to LHCI-7, respectively (Fig. 1A). These gene names are consistent with those established in previous studies (27, 44, 45). Notably, RedCAP is distinct from the LHC protein family, although it is grouped within the LHC protein superfamily (6, 7). Each LHCI subunit was identified on the basis of their characteristic amino acid residues on the cryo-EM map, specifically: L137/V139/I141 in LHCI-1, I129/F131/S133 in LHCI-2, F143/L145/A147 in LHCI-3, L132/L134/L136 in LHCI-4, F137/F139/G141 in LHCI-5, L212/F214/I216 in LHCI-6, and V144/W146/S148 in LHCI-7 (fig. S9, A to H). In addition, *G. sulphuraria* has two copies of Lhcr1 and Lhcr3 (fig. S9, I and J), among which the two Lhcr1s differ only in one residue, and the two Lhcr3s have completely the same sequence, so these have no impact on model building in this study. The root mean square deviations (RMSDs) between the structures of LHCI-4 and the other six LHCI s range from 0.99 to 3.35 Å, with LHCI-1 exhibiting a notably higher RMSD compared to the other LHCI s, reflecting its structural distinctiveness (table S4).

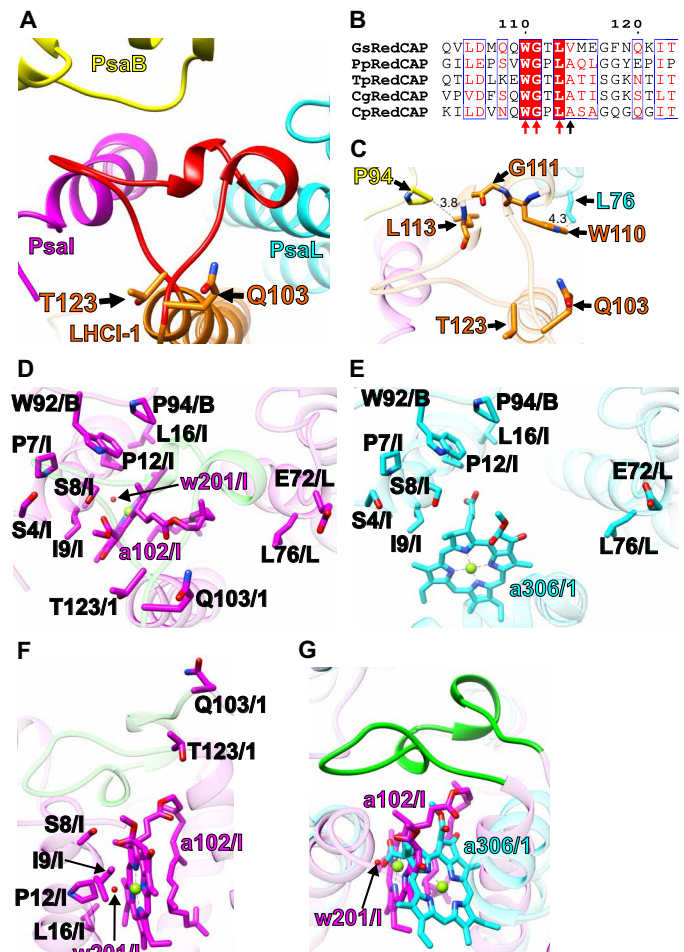
Each LHCI subunit contains several Chl and Car molecules, namely, 8 Chls *a*/6 ZXTs/1 BCX in LHCI-1, 10 Chls *a*/3 ZXTs in LHCI-2, 6 Chls *a*/1 ZXT in LHCI-3, 11 Chls *a*/3 ZXTs/1 BCR/1 BCX in LHCI-4, 11 Chls *a*/3 ZXTs/1 BCR/1 BCX in LHCI-5, 12 Chls *a*/4 ZXTs/1 BCX in LHCI-6, and 11 Chls *a*/5 ZXTs in LHCI-7 (fig. S10 and table S3). The axial ligands for the central Mg atoms of Chls in each LHCI subunit are provided by amino acid residues, as well as water and lipid molecules (table S5). Potential excitation-energy-transfer pathways from the antenna subunits to the PSI core were proposed on the basis of close physical interactions among Chls, specifically between LHCI-1 and PsaI (a), LHCI-2 and PsaB (b), LHCI-3 and PsaB (c), LHCI-4 and Psa28/PsaF (d/e), LHCI-5 and PsaJ (f), LHCI-6 and PsaA (g), and LHCI-7 and PsaK (h) (fig. S11).

Among these LHCI subunits, LHCI-6 is the gene product of *Lhcr1*. Our previous study showed that the protein band of Lhcr1 in SDS-polyacrylamide gel electrophoresis appeared at a position with a molecular weight near 29.0 kDa, which is higher than the typical molecular weight range for LHC proteins (14 to 20 kDa) (27). The GsPSI-LHCI structure shows that the LHCI-6 subunit starts from K119 (fig. S12), and its structure with three transmembrane helices is virtually identical to the other six LHCI structures (fig. S10). These findings suggest that the N-terminal extension up to K119 is disordered in the GsPSI-LHCI structure and, hence, not visible in the cryo-EM map.

### Structural characteristics of LHCI-1 (RedCAP)

The *G. sulphuraria* LHCI-1 was identified as RedCAP, a protein with unique structural features distinct from the LHC protein family (6, 7). RedCAP has been found at the LHCI-1 site in the PSI-LHCI structures of various red-lineage algae (34). The characteristic structure of RedCAP is largely defined by a loop region inserted into a cavity formed by PsaB, PsaI, and PsaL at the luminal side, as seen in the PSI-FCPI structure of the diatom *Thalassiosira pseudonana* (34). This distinctive loop, spanning residues Q103-T123 (Fig. 4A), is also present

in the GsPSI-LHCI structure. Multiple sequence alignments of RedCAP show conserved Trp, Gly, and Leu residues (red arrows in Fig. 4B and fig. S13), which are part of the Q103-T123 loop in the *G. sulphuraria* RedCAP. We recently hypothesized that the protein motif (xWGxLxx) plays a key role in RedCAP binding to PSI in the red-lineage algae (34); however, the *G. sulphuraria* RedCAP has



**Fig. 4. Structural characteristics of RedCAP.** (A) Protein arrangement of LHCI-1 (RedCAP; orange) with PsaB (yellow), PsaI (magenta), and PsaL (cyan) viewed from the luminal side. The Q103-T123 loop of RedCAP is highlighted by red. (B) Multiple sequence alignment of RedCAP of *G. sulphuraria* NIES-3638 (GsRedCAP) with those of *P. purpureum* UTEX 2757 (PpRedCAP), *T. pseudonana* CCMP1335 (TpRedCAP), *C. gracilis* UTEX LB 2658 (CgRedCAP), and *C. placodea* T11 (CpRedCAP) using ClustalW ([www.genome.jp/tools-bin/clustalw](http://www.genome.jp/tools-bin/clustalw)) and ESPript (<https://esprict.ibcp.fr/ESPript/cgi-bin/ESPript.cgi>). The Q103-T123 region was extracted from the whole alignment (fig. S13). Characteristic amino acid residues in the Q103-T123 loop are indicated by red and black arrows. (C) Protein-protein interactions of LHCI-1 with PsaL and PsaB. Interactions are indicated by dashed lines, and the numbers are distances in angstrom. Amino acid residues participating in the interactions are labeled. Each color corresponds to that in (A). (D to G) Arrangements of proteins and cofactors around the Q103-T123 loop in the GsPSI-LHCI structure [magenta in (D), (F), and (G)] and around the corresponding region in the *C. caldarium* PSI-LHCI structures (PDB: 8WEY) [cyan in (E) and (G)]. Superposition of the PSI-LHCI structures between *G. sulphuraria* and *C. caldarium* (G). The Q103-T123 loop is highlighted by green in each panel. Amino acid residues and cofactors are labeled; for example, W92/B means Trp<sup>92</sup> of PsaB. B, PsaB; I, PsaI; L, PsaL; 1, LHCI-1. The Chl and water molecules are labeled: a102/I for Chl102 of PsaI, w201/I for water201 of PsaI, and a306/1 for Chl306 of LHCI-1, which are highlighted by magenta or cyan in each PSI-LHCI.



Val instead of Ala in this loop (a black arrow in Fig. 4B), suggesting that a modified motif (xWGxLxxx) may still support the association of RedCAP with PSI (Fig. 4C).

It is noteworthy that RedCAP is present in *G. sulphuraria* NIES-3638 but absent in *C. caldarium* NIES-2137 and *C. merolae* (20–22). These species belong to the monophyletic Cyanidiophyceae (6, 7, 19, 35) distinct from Rhodophytina (19). In *C. caldarium* NIES-2137 and *C. merolae*, the LHCI-1 site is occupied by Lhcr1 (21, 22). We compared the structural environment surrounding the Q103-T123 loop of LHCI-1 in the GsPSI-LHCI with the corresponding region in *C. caldarium* (Fig. 4, D and E). The protein conformations of Psal, Psal, and Psal around the Q103-T123 loop (green cartoon) in the GsPSI-LHCI (Fig. 4D) are virtually identical with the corresponding protein conformations in the *C. caldarium* PSI-LHCI (Fig. 4E), with conserved amino acid residues W92/P94 of Psal, S4/P7/S8/I9/P12/L16 of Psal, and E72/L76 of Psal (Fig. 4, D and E). Sequence similarities (identities) for Psal, Psal, and Psal between *G. sulphuraria* and *C. caldarium* are 92% (84%), 75% (53%), and 80% (63%), respectively (fig. S14).

Substantial differences in Chl composition exist around the Q103-T123 loop. The GsPSI-LHCI contains Chl a102 of Psal (a102/I in Fig. 4D), while the *C. caldarium* PSI-LHCI has Chl a306 of LHCI-1 (a306/I in Fig. 4E). The *G. sulphuraria* and *C. caldarium* PSI-LHCI structures lack Chl molecules corresponding to the *C. caldarium* a306/I and the *G. sulphuraria* a102/I, respectively. In the GsPSI-LHCI, a water molecule w201 of Psal coordinates with the Mg atom of a102/I (Fig. 4, D and F). Despite steric hindrance caused by the two Chls (Fig. 4G), the protein conformation of Psal, especially residues S8, I9, P12, and L16, remains conserved between the two species (Fig. 4, D to F, and fig. S14). These observations suggest that the absence of RedCAP in the PSI core of *C. caldarium* NIES-2137 may lead to two changes: (i) spatial constraints imposed by PSI proteins around the Q103-T123 loop of RedCAP and/or (ii) loss of a102/I with the addition of a306/I.

We propose that during the evolutionary history of red-lineage algae, RedCAP was present in an ancestral red alga, retained by *G. sulphuraria* and *P. purpureum*, but lost in *C. caldarium* and *C. merolae* (22). Considering the characteristics of its PSI-LHCI structure, *G. sulphuraria* appears to share a close relationship with an ancestral red alga, although it belongs to the monophyletic Cyanidiophyceae (6, 7, 19, 35).

### Protein-protein interactions of other LHCI subunits

LHCI-2 (Lhcr6) is positioned near Psal through protein-protein interactions at distances of 3.3 to 3.5 Å at the stromal side, where residues N73/S145 of LHCI-2 interact with K160 of Psal (Fig. 5B). While LHCI-2 has an Asn at position 73, the other LHCI-1s have Asp at the corresponding position (fig. S15). In addition, LHCI-2 interacts with LHCI-3 through protein-protein interactions at distances of 3.5 to 4.1 Å at the stromal side between F46/L47 of LHCI-2 and I152/A156/F170 of LHCI-3 (Fig. 5C). Notably, F46 is conserved across the six LHCI-1s, whereas L47 of LHCI-2 is replaced by Leu, Met, or Phe in the other LHCI-1s (fig. S15). These results suggest that N73 of LHCI-2 plays a key role in recognizing and preferentially binding to Psal over LHCI-3.

LHCI-3 (Lhcr3) is positioned near Psal and Psal28 through interactions between residues I154/K158 of LHCI-3 and W295/S296 of Psal and between I85 of LHCI-3 and V70/P71 of Psal28, with distances of 3.5 to 4.3 Å at the stromal side (Fig. 5, C and D). Unique to LHCI-3, positions 85, 154, and 158 contain Ile, Ile, and Lys, respectively, which

are absent in the other LHCI-1s (fig. S15). This unique combination likely facilitates the specific recognition and preferential binding of I85 and I154/K158 to Psal28 and Psal, respectively.

LHCI-4 (Lhcr5) is positioned near Psal, interacting through residues L69/F71/D73 of LHCI-4 with L166/L169/R170/T173 of Psal at distances of 2.7 to 4.3 Å at the stromal side (Fig. 5E). LHCI-4 uniquely has Phe and Asp at positions 71 and 73 (fig. S15), suggesting that this distinctive combination enables selective recognition of the binding site on Psal.

LHCI-5 (Lhcr4) is also positioned near Psal, where residues M152/T153/P154/L155 of LHCI-5 interact with F157/S159/S160/F162/I163 of Psal at distances of 3.7 to 4.4 Å at the stromal side (Fig. 5E). The Met, Thr, and Pro residues at positions 152, 153, and 154 in LHCI-5 are unique among the six LHCI-1s (fig. S15), supporting its preferential binding to Psal. Furthermore, LHCI-5 interacts with LHCI-6 through protein-protein interactions between P53/F54/E56/L74/Y75 of LHCI-5 and F224/Q228/K233/F239 of LHCI-6 at distances of 2.7 to 3.8 Å at the stromal side (Fig. 5F). However, these residues are not unique to LHCI-5 (fig. S15), suggesting that M152/T153/P154 of LHCI-5 are critical for its preferential binding to Psal rather than LHCI-6.

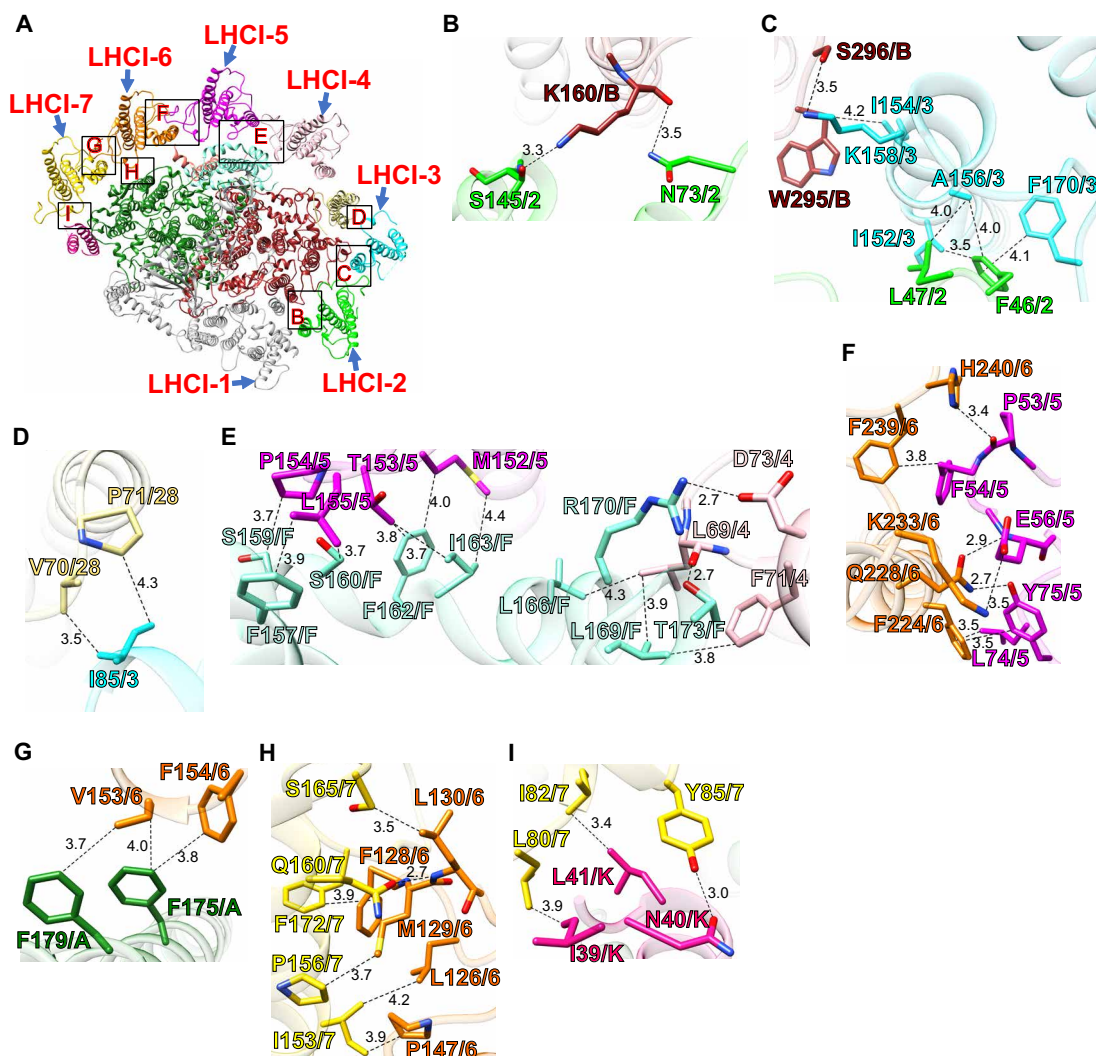
LHCI-6 (Lhcr1) is positioned near Psal through interactions between residues V153/F154 of LHCI-6 and F175/F179 of Psal at distances of 3.7 to 4.0 Å at the stromal side (Fig. 5G). LHCI-6 contains a unique combination of Val and Phe at these positions (fig. S15), facilitating specific binding to Psal. LHCI-6 also interacts with LHCI-7 via residues L126/F128/M129/L130/P147 of LHCI-6 and I153/P156/Q160/S165/F172 of LHCI-7 at distances of 3.5 to 4.2 Å at the stromal side (Fig. 5H). LHCI-6 has Met, Leu, Phe, and Lys at positions 129, 130, 224, and 233, respectively, a unique combination among the six LHCI-1s (fig. S15). This distinct set of residues (M129/L130/F224/K233) likely plays an important role in the recognition and preferential binding of LHCI-6 to Psal, LHCI-5, and LHCI-7.

LHCI-7 (Lhcr2) is positioned near Psal, interacting through residues L80/I82/Y85 of LHCI-7 with I39/N40/L41 of Psal at distances of 3.0 to 3.9 Å at the stromal side (Fig. 5I). LHCI-7 contains Tyr, Pro, and Ser at positions 85, 156, and 165, respectively, which are unique among the six LHCI-1s (fig. S15), suggesting that Y85/P156/S165 of LHCI-7 are crucial for recognizing and preferentially binding to Psal and LHCI-6.

### DISCUSSION

Kumazawa and Ifuku (25) have reported a comprehensive phylogenetic analysis of red algal LHCI-1s, classifying red algal Lhcrs into groups I to VII and a Galdieriales-specific Lhcr clade. Among the six Lhcrs of *G. sulphuraria* (GsLhcr1 to GsLhcr6), GsLhcr6, GsLhcr5, GsLhcr4, GsLhcr1, and GsLhcr2 were categorized into groups I, IV, V, VI, and VII, respectively, and corresponded to their orthologs in *P. purpureum*, namely, PpLhcr2 (group I), PpLhcr1 (group IV), PpLhcr5 (group V), PpLhcr4 (group VI), and PpLhcr3 (group VII). In addition, GsLhcr4, GsLhcr1, and GsLhcr2 were identified as the orthologs of *C. merolae* Lhcr1 (CmLhcr1), CmLhcr2, and CmLhcr3, respectively, while GsLhcr3 was uniquely classified into the Galdieriales-specific clade, with no orthologs in *P. purpureum* or *C. merolae*.

On the basis of the phylogenetic analysis (25), structural comparisons and orthologous relationships among red algal LHCI-1s are summarized in Fig. 6. According to Kumazawa and Ifuku (25), the PSI-LHCI structures of *C. merolae*, *P. purpureum*, and *G. sulphuraria*



**Fig. 5. Protein-protein interactions between LHCI and PSI subunits and among LHCIs.** (A) The PSI-LHCI structure is viewed from the stromal side. The areas encircled by black squares are enlarged in (B) to (I). The PSI and LHCI subunits involved in protein-protein interactions are colored differently. (B to I) Characteristic amino acid residues involved in protein-protein interactions. Interactions are indicated by dashed lines, and the numbers are distances in angstrom. Amino acid residues participating in the interactions are labeled; for example, K160/B and N73/2 indicate Lys<sup>160</sup> of PsaB and Asn<sup>73</sup> of LHCI-2, respectively. A, PsaA; B, PsaB; F, PsaF; K, PsaK; 28, Psa28; 2, LHCI-2; 3, LHCI-3; 4, LHCI-4; 5, LHCI-5; 6, LHCI-6; 7, LHCI-7.

are used, with *C. caldarium* RK-1 (NIES-2137) exhibiting structural consistency with *C. merolae* (22). Binding positions are assigned as p1 to p8 according to the PSI-LHCI of *P. purpureum* (PpPSI-LHCI), which has the largest number of LHCI subunits among the three species. In GsPSI-LHCI, the binding sites of LHCI subunits correspond to LHCI-1 (p1), LHCI-2 (p2), LHCI-3 (p3), LHCI-4 (p5), LHCI-5 (p6), LHCI-6 (p7), and LHCI-7 (p8).

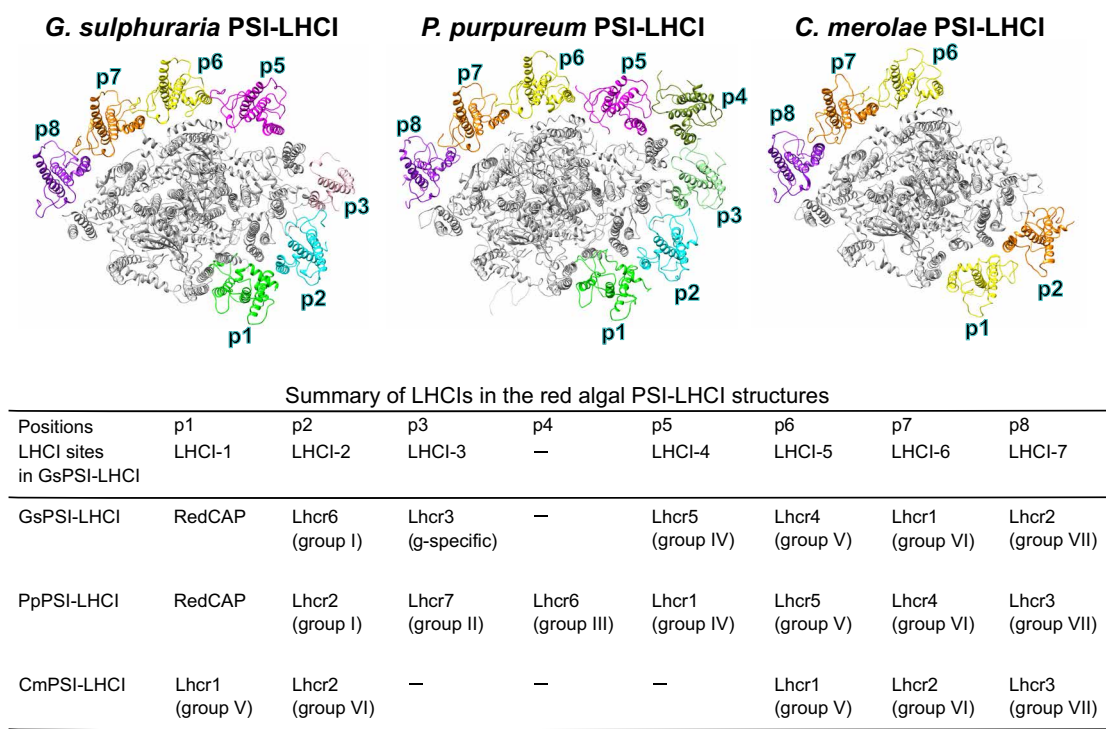
The GsLhcrs at the LHCI-2 and LHCI-4 to LHCI-7 sites in GsPSI-LHCI correspond to their orthologs at the respective sites of PpPSI-LHCI, whereas GsLhcr3 (a Galdieriales-specific Lhcr clade) at the LHCI-3 site in GsPSI-LHCI aligns with PpLhcr7 (group II) at the corresponding site in PpPSI-LHCI, despite their lack of orthology. In contrast, the GsLhcrs at the LHCI-5 to LHCI-7 sites in GsPSI-LHCI correspond to orthologs of the PSI-LHCI of *C. merolae* (CmPSI-LHCI) at the respective sites, although GsLhcr6 at the LHCI-2 site in GsPSI-LHCI is replaced by CmLhcr2 (group VI) at the corresponding site in CmPSI-LHCI. These observations indicate that the group

VI Lhcr at the LHCI-2 site in CmPSI-LHCI has been acquired secondarily after the divergence of Galdieriales from other Cyanidiophyceae including Cyanidiales and Cyanidioschyzonales, supporting neolocalization of this LHCI based on combined structural and phylogenetic analysis (22).

We compared amino acid residues of LHCIs among the three red algal PSI-LHCI structures (fig. S16) based on the protein-protein interaction data from GsPSI-LHCI (Figs. 4 and 5) and the summary of structural comparisons and orthologous relationships among LHCI subunits (Fig. 6). The amino acid residues involved in protein-protein interactions between LHCI and PSI subunits are largely conserved between *G. sulphuraria* and *C. merolae*/P. *purpureum* at the LHCI-4, LHCI-6, and LHCI-7 sites (fig. S16, B, D, and E), whereas they differ at the LHCI-2 and LHCI-5 sites (fig. S16, A and C).

It is of particular interest that the binding properties of fucoxanthin Chl *a/c*-binding proteins (FCPs) in the PSI-FCPI structures between *C. gracilis* and *T. pseudonana* are highly conserved (34). The





**Fig. 6. Structural and phylogenetic diversity of red algal LHCIs.** The PSI-LHCI structures of the three red algae, *G. sulphuraria* (this study), *P. purpureum* (PDB: 7Y5E), and *C. merolae* (PDB: 5ZGB), are viewed from the stromal side. Each LHCI position is labeled p1 to p8 based on the *P. purpureum* PSI-LHCI. PSI subunits are shown in gray, whereas LHCI subunits are depicted in various other colors. LHCI subunits belonging to the same group are shown in the same color. The table summarizes the correlations between binding positions (p1 to p8) and each of gene products identified, LHCI sites in the GsPSI-LHCI, and groups defined previously (25).

orthologous relationship and amino acid residues involved in interactions closely match between the two diatoms. In contrast, among the three red algae, the conservation of amino acid residues involved in protein-protein interactions is significantly low, despite clear orthologous relationship among LHCIs. These observations imply that protein-protein interactions between LHCI and PSI subunits were less conserved during the early evolutionary stages of red algae but became highly conserved following secondary endosymbiosis, as seen in diatoms.

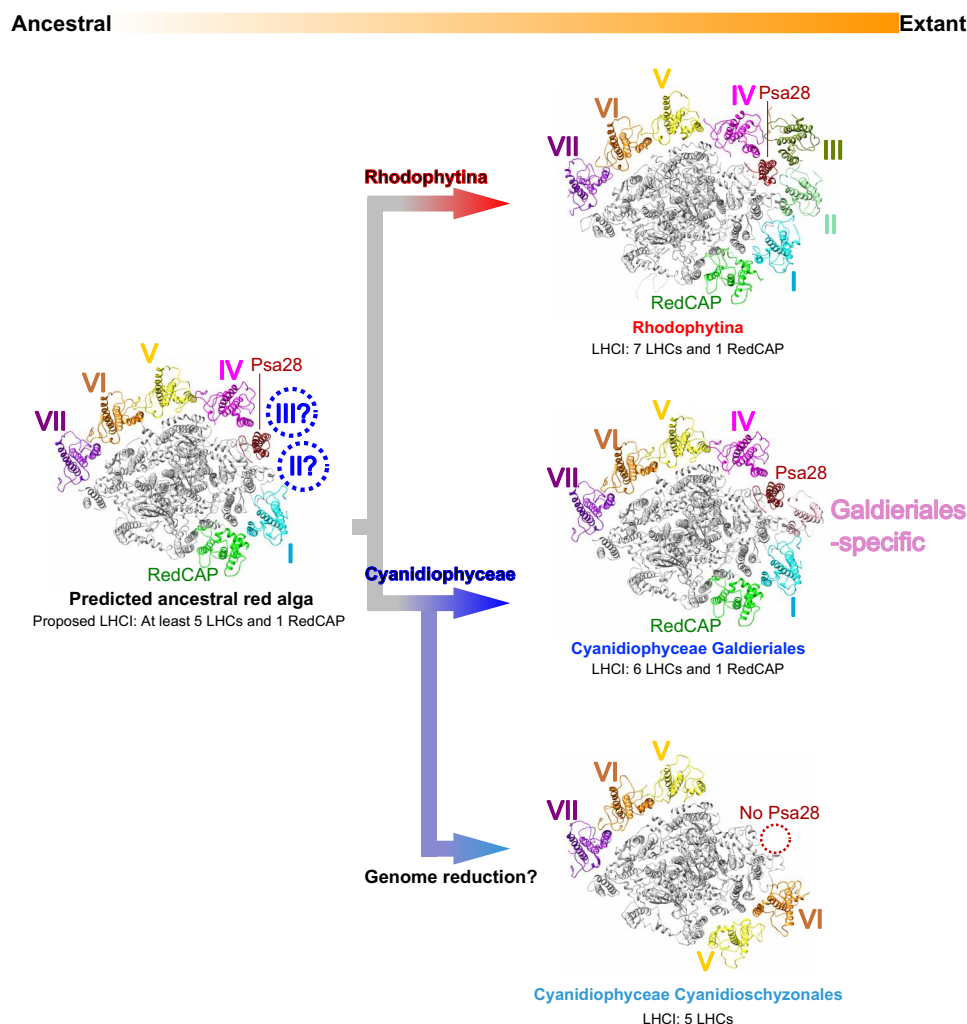
This study has markedly advanced our understanding of the evolutionary trajectory of PSI-LHCI in red algae by elucidating the structure of PSI-LHCI from *G. sulphuraria*, a member of the Galdieriales—the earliest diverging group within Cyanidiophyceae. Leveraging these findings, we propose a model for an ancestral red algal PSI-LHCI supercomplex (Fig. 7), which likely features RedCAP and Lhcrs from groups I and IV to VII in conserved positions across both Rhodophytina and Galdieriales, although these components are absent in Cyanidioschyzonales. It is noteworthy that LHCIs in Cyanidioschyzonales lack ancestral traits at positions 1 and 2 due to neolocalization, reflecting LHCI reorganization during genome reduction (22). In addition, we hypothesize that Psa28 was present in an ancestral red algal PSI-LHCI, fulfilling a pivotal role in group IV Lhcr binding.

The evolutionary trajectory of LHCI subunits at positions 3 and 4, as shown in Fig. 6, permits us to propose plausible scenarios (fig. S17). For position 3, it remains conceivable that either the group II Lhcr in PpPSI-LHCI or the Galdieriales-specific Lhcr in GsPSI-LHCI originated from a common ancestor. Alternatively, both may have

arisen independently after divergence from a common ancestor. The close interactions between group II and group III Lhcrs within Rhodophytina PSI-LHCI suggest that these Lhcrs were acquired concurrently, irrespective of the exact timing of acquisition. This results in six equally plausible evolutionary scenarios based on the current Lhcr phylogenetic topology. Consequently, reconstructing the complete architecture of an ancestral red algal PSI-LHCI remains a challenge. Future research integrating both structural and phylogenetic analyses of red-lineage LHCs may offer critical insights into the binding mechanisms of LHCI subunits, including the identification of secondarily acquired Lhcrs. This approach could potentially aid in reconstructing the ancestral PSI-LHCI structure.

MATERIALS AND METHODS  
HPLC analysis

Culture of the *G. sulphuraria* NIES-3638 cells from the Microbial Culture Collection at National Institute for Environmental Studies (NIES), Japan, and preparation of the PSI-LHCI supercomplex were performed, as described previously (27). Quinone molecules were analyzed using a Shimadzu HPLC system (46) equipped with a reversed-phase InertSustain C18 column (250 mm by 4.6 mm, 5-μm particle size; GL Sciences). Quinone and pigment molecules in the prepared *G. sulphuraria* PSI cores (27) were extracted using 100% methanol. The samples were subjected to HPLC and eluted with solvent A (100% methanol) and solvent B (100% isopropanol) in a linear gradient: 0 to 10 min, 0% solvent B; 10 to 30 min, 0 to 97% solvent B; and 30 to 40 min, 97% solvent B. The flow rate was set to 0.5 ml



**Fig. 7. Evolutionary scheme of LHCI subunits in red algal PSI-LHCI supercomplexes.** The PSI-LHCI structures are viewed from the stromal side. LHCI and Psa28 subunits are colored, while other PSI subunits are shown in gray. LHCI subunits belonging to the same group are shown in the same color. I to VII, groups I to VII of Lhcrcs. An ancestral red algal PSI-LHCI structure is provisionally depicted using the GsPSI-LHCI structure, whereas the other PSI-LHCI structures originate from the Rhodophytina *P. purpureum* (PDB: 7Y5E), the Galdieriales *G. sulphuraria* (this study), and the Cyanidioschyzonales *C. merolae* (PDB: 5ZGB).

$\text{min}^{-1}$ , with detection at 270 nm. Phylloquinone (P0642, Tokyo Chemical Industry) and ubiquinone-4 (C2470, Sigma) were purchased as standards.

### Cryo-EM data collection

A 3- $\mu\text{l}$  aliquot of the GsPSI-LHCI supercomplex (2.5 mg of Chl  $\text{ml}^{-1}$ ) in 20 mM Mes-NaOH (pH 6.5) buffer containing 0.5 M betaine, 5 mM  $\text{CaCl}_2$ , 10 mM  $\text{MgCl}_2$ , and 0.03% *n*-dodecyl- $\beta$ -D-maltoside was applied to a Quantifoil R1.2/1.3 Cu 300 mesh grid in the chamber of FEI Vitrobot Mark IV (Thermo Fisher Scientific). Then, the grid was blotted with a filter paper for 4 s at 4°C under 100% humidity and plunged into liquid ethane cooled by liquid nitrogen. The frozen grid was transferred into a CRYO ARM 300 electron microscope (JEOL) equipped with a cold-field emission gun operated at 300 kV. All image stacks were collected from  $5 \times 5$  holes per stage adjustment to the central hole, and image shifts were applied to the surrounding holes while maintaining an axial coma-free condition. The images were recorded with an in-column energy filter with a slit width of 20 eV and

at a nominal magnification of  $\times 60,000$  on a direct electron detector (Gatan K3, AMETEK). The nominal defocus range was  $-1.8$  to  $-1.2 \mu\text{m}$ . Physical pixel size corresponded to  $0.752 \text{ \AA}$ . Each image stack was exposed at a dose rate of  $25.34 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$  for 1.97 s in correlated double sampling (CDS) mode, with dose-fractionated 50 movie frames. In total, 8800 image stacks were collected.

### Cryo-EM image processing

The resultant movie frames were aligned and summed using MotionCor2 (47) to yield dose-weighted images. Estimation of the contrast transfer function was performed using CTFFIND4 (48). All of the following processes were performed using RELION-4.0 (49). In total, 2,583,694 particles were automatically picked up and used for reference-free two-dimensional (2D) classification. Then, 933,316 particles were selected from good 2D classes and subsequently subjected to 3D classification without any symmetry. An initial model for the first 3D classification was generated de novo from 2D classification. As shown in fig. S1C, the final PSI-LHCI structure was

reconstructed from 110,313 particles. The overall resolution of the cryo-EM map was estimated to be 2.19 Å by the gold standard Fourier Shell Correlation (FSC) curve with a cutoff value of 0.143 (fig. S2A) (50). Local resolutions were calculated using RELION (fig. S2C).

## Model building and refinement

Two types of the cryo-EM maps were used for the model building of the PSI-LHCI supercomplex: One was a postprocessed map, and the other was a denoised map using Topaz version 0.2.4 (51). The postprocessed map was denoised using the trained model in 100 epochs with two half-maps. Homology models of each subunit in the PSI-LHCI supercomplex were built using the Phyre2 server (52), and then their structures were inspected and manually adjusted against the maps with Coot (53). Each model was built on the basis of interpretable features from the density maps at a contour level of 1.5  $\sigma$  in the denoised and postprocessed maps. For the assignment of Cars, BCR, BCX, and ZXT were distinguished on the basis of the density covering the head group of Cars with the above thresholds. The PSI-LHCI structure was refined with phenix.real\_space\_refine (54) and Servalcat (55), with geometric restraints for the protein-cofactor coordination. The final model was validated with MolProbity (56), EMRinger (57), and Q-score (58). The statistics for all data collection and structure refinement are summarized in tables S1 and S2. All structural figures were made by PyMOL (59), UCSF Chimera (60), and UCSF ChimeraX (61).

Because the numbering of Chls, Cars, and other cofactors in this paper was different from those of the Protein Data Bank (PDB) data, we listed the relationship of their numbering in this paper with those in the PDB data in tables S6 to S8.

## Phylogenetic analysis

Psa28 sequences were collected from National Center for Biotechnology Information (NCBI) nr, NCBI tsa-nr, Joint Genome Institute (JGI) Phycocosm, and EukProt v3 (<https://evocellbio.com/eukprot/>) (62), and nuclear draft genome assemblies were deposited in a public repository using BLASTP (63). Accession numbers were provided on the molecular phylogenetic tree of Psa28. The Psa28 sequences of *Diacronema viridis*, *Phaeocystis jahnii*, and *Phaeocystis rex* are derived from their draft genomes (63). Amino acid sequences of Psa28 were manually curated and aligned using MAFFT E-INS-i v7.525 (64). The alignment was trimmed using ClipKit v2.2.0 with smart-gap mode (64). The phylogenetic tree was inferred using IQ-TREE 2.2.2.7 (65) with the Qpfam+F+R5 model (66) selected by ModelFinder (67). The tree was visualized by iTOL v6 (68) and rooted with the Galdieriales clade. Ultrafast bootstrap approximation was performed with 1000 replicates (66). SH-like approximate likelihood ratio test (SH-aLRT) support (%) tests with 1000 replicates and aBayes support tests were also performed (69, 70).

## Supplementary Materials

This PDF file includes:

Figs. S1 to S17

Tables S1 to S8

## REFERENCES AND NOTES

- R. E. Blankenship, *Molecular Mechanisms of Photosynthesis* (Wiley-Blackwell, ed. 3, 2021).
- J. H. Golbeck, Structure and function of photosystem I. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **43**, 293–324 (1992).
- K. Brettel, W. Leibl, Electron transfer in photosystem I. *Biochim. Biophys. Acta Bioenerg.* **1507**, 100–114 (2001).
- J.-R. Shen, The structure of photosystem II and the mechanism of water oxidation in photosynthesis. *Annu. Rev. Plant Biol.* **66**, 23–48 (2015).
- D. Shevela, J. F. Kern, G. Govindjee, J. Messinger, Solar energy conversion by photosystem II: Principles and structures. *Photosynth. Res.* **156**, 279–307 (2023).
- J. Engelken, H. Brinkmann, I. Adamska, Taxonomic distribution and origins of the extended LHC (light-harvesting complex) antenna protein superfamily. *BMC Evol. Biol.* **10**, 233 (2010).
- S. Sturm, J. Engelken, A. Gruber, S. Vugrinec, P. G. Kroth, I. Adamska, J. Lavaud, A novel type of light-harvesting antenna protein of red algal origin in algae with secondary plastids. *BMC Evol. Biol.* **13**, 159 (2013).
- P. G. Falkowski, M. E. Katz, A. H. Knoll, A. Quigg, J. A. Raven, O. Schofield, F. J. R. Taylor, The evolution of modern eukaryotic phytoplankton. *Science* **305**, 354–360 (2004).
- M. Hippler, N. Nelson, The plasticity of photosystem I. *Plant Cell Physiol.* **62**, 1073–1081 (2021).
- J.-R. Shen, "Structure, function, and variations of the photosystem I-antenna supercomplex from different photosynthetic organisms," in *Macromolecular Protein Complexes IV. Subcellular Biochemistry*, J. R. Harris, J. Marles-Wright, Eds. (Springer, 2022), vol. 99, pp. 351–377.
- H. S. Yoon, J. D. Hackett, C. Ciniglia, G. Pinto, D. Bhattacharya, A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* **21**, 809–818 (2004).
- B. R. Green, E. Pichersky, Hypothesis for the evolution of three-helix Chl a/b and Chl a/c light-harvesting antenna proteins from two-helix and four-helix ancestors. *Photosynth. Res.* **39**, 149–162 (1994).
- B. R. Green, D. G. Durnford, The chlorophyll-carotenoid proteins of oxygenic photosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 685–714 (1996).
- C. Büchel, Light harvesting complexes in chlorophyll c-containing algae. *Biochim. Biophys. Acta Bioenerg.* **1861**, 148027 (2020).
- H. S. Yoon, G. C. Zuccarello, D. Bhattacharya, in *Red Algae in the Genomic Age*, J. Seckbach, D. J. Chapman, Eds. (Springer, 2010), pp. 25–42.
- F. D. Ott, J. Seckbach, in *Evolutionary Pathways and Enigmatic Algae: Cyanidium caldarium (Rhodophyta) and Related Cells*, J. Seckbach, Ed. (Springer, 1994), vol. 91, pp. 145–152.
- F. D. Ott, *Handbook of the Taxonomic Names Associated with the Non-marine Rhodophycophyta* (J. Cramer, 2009).
- S.-L. Liu, Y.-R. Chiang, H. S. Yoon, H.-Y. Fu, Comparative genome analysis reveals *Cyanidiococcus* gen. nov., A new extremophilic red algal genus sister to *Cyanidioschyzon* (Cyanidioschyzonaceae, Rhodophyta). *J. Phycol.* **56**, 1428–1442 (2020).
- S. I. Park, C. H. Cho, C. Ciniglia, T.-Y. Huang, S.-L. Liu, D. E. Bustamante, M. S. Calderon, A. Mansilla, T. McDermott, R. A. Andersen, H. S. Yoon, Revised classification of the Cyanidiophyceae based on plastid genome data with descriptions of the Cavemulicolales ord. nov. and Galdieriales ord. nov. (Rhodophyta). *J. Phycol.* **59**, 444–466 (2023).
- M. Antoshvili, I. Caspy, M. Hippler, N. Nelson, Structure and function of photosystem I in *Cyanidioschyzon merolae*. *Photosynth. Res.* **139**, 499–508 (2019).
- X. Pi, L. Tian, H.-E. Dai, X. Qin, L. Cheng, T. Kuang, S.-F. Sui, J.-R. Shen, Unique organization of photosystem I-light-harvesting supercomplex revealed by cryo-EM from a red alga. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4423–4428 (2018).
- K. Kato, T. Hamaguchi, M. Kumazawa, Y. Nakajima, K. Ifuku, S. Hirooka, Y. Hirose, S.-y. Miyagishima, T. Suzuki, K. Kawakami, N. Dohmae, K. Yonekura, J.-R. Shen, R. Nagao, The structure of PSI-LHCI from *Cyanidium caldarium* provides evolutionary insights into conservation and diversity of red-lineage LHCS. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2319658121 (2024).
- X. You, X. Zhang, J. Cheng, Y. Xiao, J. Ma, S. Sun, X. Zhang, H.-W. Wang, S.-F. Sui, In situ structure of the red algal phycobilisome-PSII-PSI-LHC megacomplex. *Nature* **616**, 199–206 (2023).
- C. H. Cho, S. I. Park, T.-Y. Huang, Y. Lee, C. Ciniglia, H. C. Yadavalli, S. W. Yang, D. Bhattacharya, H. S. Yoon, Genome-wide signatures of adaptation to extreme environments in red algae. *Nat. Commun.* **14**, 10 (2023).
- M. Kumazawa, K. Ifuku, Unraveling the evolutionary trajectory of LHCI in red-lineage algae: Conservation, diversification, and neolocalization. *iScience* **27**, 110897 (2024).
- B. Thangaraj, C. C. Jolley, I. Sarrou, J. B. Bultema, J. Greyslak, J. P. Whitelegge, S. Lin, R. Kouřil, R. Subramanyam, E. J. Boekema, P. Fromme, Efficient light harvesting in a dark, hot, acidic environment: The structure and function of PSI-LHCI from *Galdieria sulphuraria*. *Biophys. J.* **100**, 135–143 (2011).
- R. Nagao, H. Ogawa, T. Suzuki, N. Dohmae, K. Kato, Y. Nakajima, J.-R. Shen, Biochemical evidence for the diversity of LHCI proteins in PSI-LHCI from the red alga *Galdieria sulphuraria* NIES-3638. *Photosynth. Res.* **163**, 14 (2025).
- R. Nagao, K. Kato, K. Ifuku, T. Suzuki, M. Kumazawa, I. Uchiyama, Y. Kashino, N. Dohmae, S. Akimoto, J.-R. Shen, N. Miyazaki, F. Akita, Structural basis for assembly and function of a diatom photosystem I-light-harvesting supercomplex. *Nat. Commun.* **11**, 2481 (2020).



29. C. Xu, X. Pi, Y. Huang, G. Han, X. Chen, X. Qin, G. Huang, S. Zhao, Y. Yang, T. Kuang, W. Wang, S.-F. Sui, J.-R. Shen, Structural basis for energy transfer in a huge diatom PSI-FCPI supercomplex. *Nat. Commun.* **11**, 5081 (2020).
30. L.-S. Zhao, P. Wang, K. Li, Q.-B. Zhang, F.-Y. He, C.-Y. Li, H.-N. Su, X.-L. Chen, L.-N. Liu, Y.-Z. Zhang, Structural basis and evolution of the photosystem I-light-harvesting supercomplex of cryptophyte algae. *Plant Cell* **35**, 2449–2463 (2023).
31. S. Zhang, L. Si, X. Su, X. Zhao, X. An, M. Li, Growth phase-dependent reorganization of cryptophyte photosystem I antennae. *Commun. Biol.* **7**, 560 (2024).
32. L.-S. Zhao, N. Wang, K. Li, C.-Y. Li, J.-P. Guo, F.-Y. He, G.-M. Liu, X.-L. Chen, J. Gao, L.-N. Liu, Y.-Z. Zhang, Architecture of symbiotic dinoflagellate photosystem I-light-harvesting supercomplex in *Symbiodinium*. *Nat. Commun.* **15**, 2392 (2024).
33. X. Li, Z. Li, F. Wang, S. Zhao, C. Xu, Z. Mao, J. Duan, Y. Feng, Y. Yang, L. Shen, G. Wang, Y. Yang, L.-J. Yu, M. Sang, G. Han, X. Wang, T. Kuang, J.-R. Shen, W. Wang, Structures and organizations of PSI-AcpPCI supercomplexes from red tidal and coral symbiotic photosynthetic dinoflagellates. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2315476121 (2024).
34. K. Kato, Y. Nakajima, J. Xing, M. Kumazawa, H. Ogawa, J.-R. Shen, K. Ifuku, R. Nagao, Structural basis for molecular assembly of fucoxanthin chlorophyll *a/c*-binding proteins in a diatom photosystem I supercomplex. *eLife* **13**, RP99858 (2024).
35. J. I. Kim, C. E. Moore, J. M. Archibald, D. Bhattacharya, G. Yi, H. S. Yoon, W. Shin, Evolutionary dynamics of cryptophyte plastid genomes. *Genome Biol. Evol.* **9**, 1859–1872 (2017).
36. S. Ohashi, T. Iemura, N. Okada, S. Itoh, H. Furukawa, M. Okuda, M. Ohnishi-Kameyama, T. Ogawa, H. Miyashita, T. Watanabe, S. Itoh, H. Oh-oka, K. Inoue, M. Kobayashi, An overview on chlorophylls and quinones in the photosystem I-type reaction centers. *Photosynth. Res.* **104**, 305–319 (2010).
37. H.-U. Schoeder, W. Lockau, Phyloquinone copurifies with the large subunit of photosystem I. *FEBS Lett.* **199**, 23–27 (1986).
38. Y. Takahashi, K. Hirota, S. Katoh, Multiple forms of P700-chlorophyll *a*-protein complexes from *Synechococcus* sp.: The iron, quinone and carotenoid contents. *Photosynth. Res.* **6**, 183–192 (1985).
39. Y. Ikeda, M. Komura, M. Watanabe, C. Minami, H. Koike, S. Itoh, Y. Kashino, K. Satoh, Photosystem I complexes associated with fucoxanthin-chlorophyll-binding proteins from a marine centric diatom, *Chaetoceros gracilis*. *Biochim. Biophys. Acta Bioenerg.* **1777**, 351–361 (2008).
40. M. Mimuro, T. Tsuchiya, H. Inoue, Y. Sakuragi, Y. Itoh, T. Gotoh, H. Miyashita, D. A. Bryant, M. Kobayashi, The secondary electron acceptor of photosystem I in *Gloeobacter violaceus* PCC 7421 is menaquinone-4 that is synthesized by a unique but unknown pathway. *FEBS Lett.* **579**, 3493–3496 (2005).
41. E. Yoshida, A. Nakamura, T. Watanabe, Reversed-phase HPLC determination of chlorophyll *a'* and naphthoquinones in photosystem I of red algae: Existence of two menaquinone-4 molecules in photosystem I of *Cyanidium caldarium*. *Anal. Sci.* **19**, 1001–1005 (2003).
42. R. Hasegawa, K. Saito, T. Takaoka, H. Ishikita,  $pK_a$  of ubiquinone, menaquinone, phyloquinone, plastoquinone, and rhodoquinone in aqueous solution. *Photosynth. Res.* **133**, 297–304 (2017).
43. P. Fromme, P. Jordan, N. Krauß, Structure of photosystem I. *Biochim. Biophys. Acta Bioenerg.* **1507**, 5–31 (2001).
44. J. Marquardt, S. Wans, E. Rhiel, A. Randolph, W. E. Krumbein, Intron-exon structure and gene copy number of a gene encoding for a membrane-intrinsic light-harvesting polypeptide of the red alga *Galdieria sulphuraria*. *Gene* **255**, 257–265 (2000).
45. J. Marquardt, B. Lutz, S. Wans, E. Rhiel, W. E. Krumbein, The gene family coding for the light-harvesting polypeptides of photosystem I of the red alga *Galdieria sulphuraria*. *Photosynth. Res.* **68**, 121–130 (2001).
46. R. Nagao, M. Yokono, Y. Ueno, J.-R. Shen, S. Akimoto, Low-energy chlorophylls in fucoxanthin chlorophyll *a/c*-binding protein conduct excitation energy transfer to photosystem I in diatoms. *J. Phys. Chem. B* **123**, 66–70 (2019).
47. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. F. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
48. J. A. Mindell, N. Grigorieff, Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
49. D. Kimanius, L. Dong, G. Sharov, T. Nakane, S. H. W. Scheres, New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021).
50. N. Grigorieff, S. C. Harrison, Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Curr. Opin. Struc. Biol.* **21**, 265–273 (2011).
51. T. Bepko, K. Kelley, A. J. Noble, B. Berger, Topaz-Denoise: General deep denoising models for cryoEM and cryoET. *Nat. Commun.* **11**, 5208 (2020).
52. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
53. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Cool. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
54. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
55. K. Yamashita, C. M. Palmer, T. Burnley, G. N. Murshudov, Cryo-EM single-particle structure refinement and map calculation using *Servalcat*. *Acta Crystallogr. D Struct. Biol.* **77**, 1282–1291 (2021).
56. V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
57. B. A. Barad, N. Echols, R. Y.-R. Wang, Y. Cheng, F. DiMaio, P. D. Adams, J. S. Fraser, EMRinger: Side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* **12**, 943–946 (2015).
58. G. Pintilie, K. Zhang, Z. Su, S. Li, M. F. Schmid, W. Chiu, Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat. Methods* **17**, 328–334 (2020).
59. L. L. C. Schrödinger, The PyMOL molecular graphics system. Version 2.5.0. (2021).
60. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
61. E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
62. A. Priyam, B. J. Woodcroft, V. Rai, I. Moghul, A. Munagala, F. Ter, H. Chowdhary, I. Pieniak, L. J. Maynard, M. A. Gibbins, H. Moon, A. Davis-Richardson, M. Uludag, N. S. Watson-Haigh, R. Challis, H. Nakamura, E. Favreau, E. A. Gómez, T. Pluskal, G. Leonard, W. Rumpf, Y. Wurm, Sequenceserver: A modern graphical user interface for custom BLAST databases. *Mol. Biol. Evol.* **36**, 2922–2924 (2019).
63. P. Yang, K. Guo, Y. Yang, M. Lyu, J. Liu, X. Li, Y. Feng, Phylogeny and genetic variations of the three genome compartments in haptophytes shed light on the rapid evolution of coccolithophores. *Gene* **887**, 147716 (2023).
64. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
65. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
66. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
67. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermin, ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
68. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
69. S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
70. M. Anisimova, M. Gil, J.-F. Dufayard, C. Dessimoz, O. Gascuel, Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* **60**, 685–699 (2011).
71. E. C. Yang, S. M. Boo, D. Bhattacharya, G. W. Saunders, A. H. Knoll, S. Fredericq, L. Graf, H. S. Yoon, Divergence time estimates and the evolution of major lineages in the florideophyte red algae. *Sci. Rep.* **6**, 21361 (2016).

**Acknowledgments:** We thank K. Kato and S. Kakiuchi for assistance in this study. Cryo-EM data were obtained using EM01CT, EM02CT, and EM04CT of SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute [JASRI proposal nos. 2022B2728 (J.-R.S.) and 2023A2715 (Y.N.)]. **Funding:** This work was supported by JSPS KAKENHI grant nos. JP22KJ2017 (M.K.), JP23K14211 (Y.N.), JP22H04916 (J.-R.S.), JP23H02347 (K.I.), and JP23H02423 (R.N.); Takeda Science Foundation (K.K.); and Research Support Project for Life Science and Drug Discovery [Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)] from AMED under grant no. JP22ama121001 (J.-R.S.). **Author contributions:** K.K.: Writing—original draft, investigation, and funding acquisition. M.K.: Writing—original draft, investigation, and funding acquisition. Y.N.: Writing—original draft, investigation, and funding acquisition. T.S.: Investigation. N.D.: Investigation. J.-R.S.: Funding acquisition, writing—review and editing, and supervision. K.I.: Writing—original draft, funding acquisition, writing—review and editing, and supervision. R.N.: Writing—original draft, conceptualization, investigation, writing—review and editing, methodology, funding acquisition, validation, supervision, project administration, visualization, and resources. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Atomic coordinates, cryo-EM maps, and raw image data for the reported structure have been deposited in the PDB under accession code 9KC5 (<https://doi.org/10.2210/pdb9KC5/pdb>), in the Electron

Microscopy Data Bank under accession code EMD-62242 ([www.ebi.ac.uk/emdb/EMD-62242](http://www.ebi.ac.uk/emdb/EMD-62242)), and in the Electron Microscopy Public Image Archive under accession code EMPIAR-12385 (<https://doi.org/10.6019/EMPIAR-12385>), respectively. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 6 January 2025  
Accepted 11 April 2025  
Published 16 May 2025  
[10.1126/sciadv.adv7488](https://doi.org/10.1126/sciadv.adv7488)