Light-dependent chlorophyll *f* synthase is a highly divergent paralog of PsbA of photosystem II

Ming-Yang Ho[1], Gaozhong Shen[1], Daniel P. Canniffe[1], Chi Zhao[1], and Donald A. Bryant[1, 2, 3*]

[1]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802 USA; [2]Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717 USA; [3]Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore.

***Corresponding author: e-mail: dab14@psu.edu**

Abstract

Chlorophyll (Chl) *f* permits some cyanobacteria to expand the spectral range for photosynthesis by absorbing far-red light. We used reverse genetics and heterologous expression to identify the enzyme for Chl *f* synthesis. Null mutants of "super-rogue" *psbA4* genes, divergent paralogs of *psbA* genes encoding the D1 core subunit of Photosystem II, abolished Chl *f* synthesis in two cyanobacteria that grow in far-red light. Heterologous expression of the *psbA4* gene, which we rename *chlF*, enables Chl *f* biosynthesis in *Synechococcus* sp. PCC 7002. Because the reaction requires light, Chl *f* synthase is probably a photo-oxidoreductase that employs catalytically useful Chl *a* molecules, tyrosine Y_Z, and plastoquinone like Photosystem II but lacks a $Mn_4Ca_1O_5$ cluster. Introduction of Chl *f* biosynthesis into crop plants could expand their solar energy utilization.

Structured Abstract

INTRODUCTION

Terrestrial cyanobacteria often occur in environments that receive strongly filtered light because of shading by plants or because of their associations with soil crusts, benthic mat communities, or dense cyanobacterial blooms. The light in such environments becomes highly enriched in far-red light (FRL) (wavelengths >700 nm). Cyanobacteria that are able to utilize FRL for photosynthesis have evolved a novel far-red light photoacclimation mechanism (FaRLiP) to gain a strong selective advantage over other cyanobacteria. The FaRLiP response involves extensive remodeling of photosystems I and II (PSI and PSII), and light-harvesting phycobilisome complexes. FaRLiP cells synthesize chlorophyll (Chl) *f,* Chl *d,* and FRL-absorbing phycobiliproteins under these conditions and thus can utilize FRL efficiently for oxygenic photosynthesis. A key element of the FaRLiP response is the FRL-specific expression of seventeen genes that encode paralogs of core components of the three light-harvesting complexes produced during growth in white light.

RATIONALE

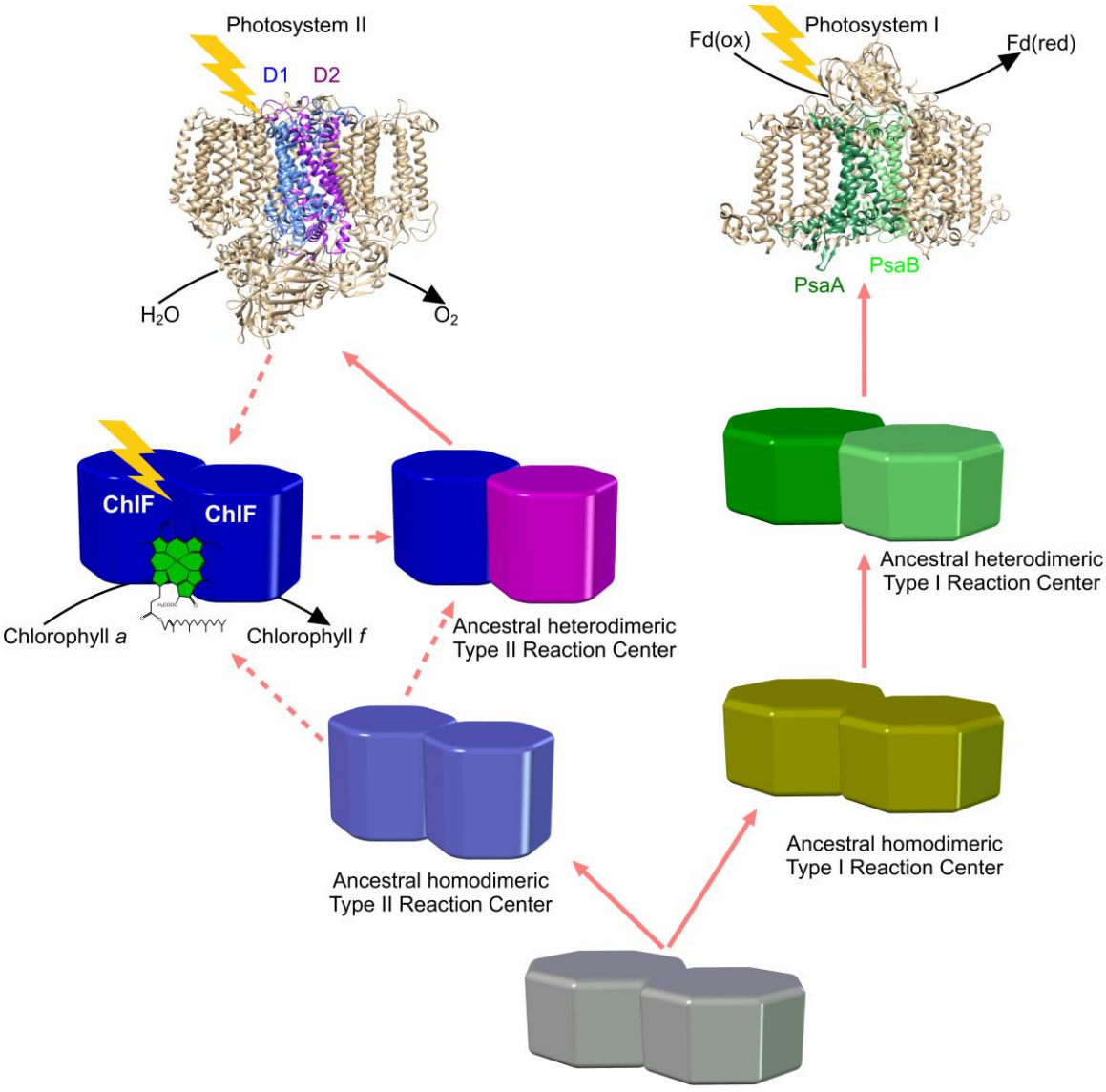
The ability to synthesize Chl *f* is a key element of the FaRLiP response, but the Chl *f* synthase was still unknown. Transcription and phylogenetic profiling suggested that the gene(s) responsible for this activity was in the conserved FaRLiP gene cluster. This led us to focus on *psbA4,* a divergent member of the *psbA* gene family encoding so-called “super-rogue” PsbA, a divergent paralog of the D1 core subunit of PSII. We used reverse genetics and heterologous expression to identify the Chl *f* synthase of two cyanobacteria capable of FaRLiP: *Chlorogloeopsis fritschii* PCC 9212 and *Synechococcus* sp. PCC 7335.

RESULTS

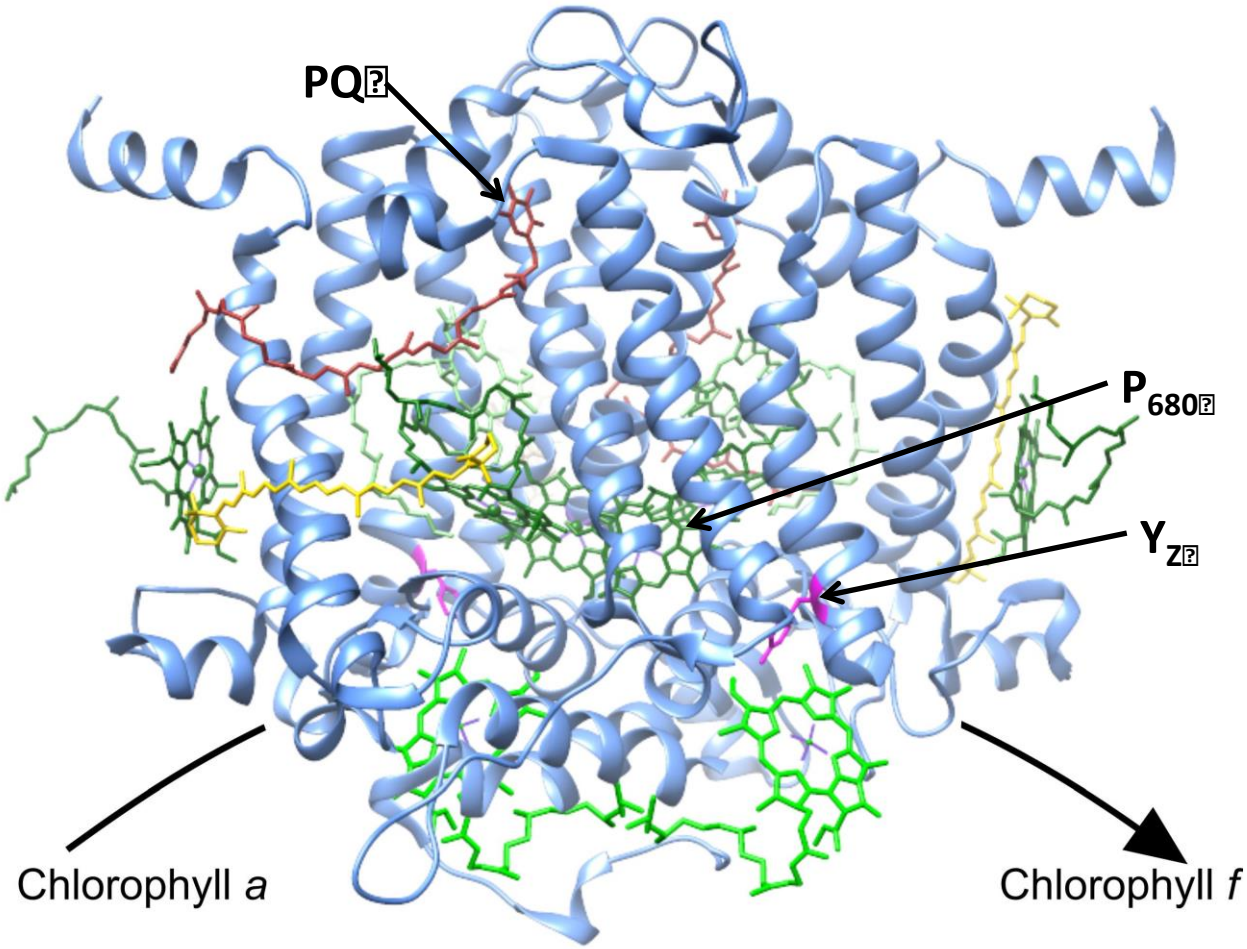
In both species, null mutants of *psbA4* no longer synthesized Chl *f* and lacked FRL absorbance and long-wavelength fluorescence emission, key spectroscopic properties associated with Chl *f* synthesis. Heterologous expression of the *psbA4* gene from *Chl. fritschii* PCC 9212 in the model, non-FaRLiP cyanobacterium *Synechococcus* sp. PCC 7002 led to the synthesis of Chl *f*. These results showed that *psbA4*, renamed ChlF, encodes the Chl *f* synthase. Growth experiments employing intervals of FRL and darkness showed that Chl *f* synthesis is light-dependent; this implied that ChlF is a photo-oxidoreductase that oxidizes Chl *a* (or Chlide *a*) instead of water.

CONCLUSIONS

ChlF may have evolved after gene duplication from PsbA of a water-oxidizing PSII complex by loss of the ligands for binding the $Mn_4Ca_1O_5$ cluster but retention of catalytically useful chlorophylls, tyrosine Y_Z, and plastoquinone binding. Alternatively, PsbA may have arisen by gene duplication from ChlF and by gaining the capacity to bind the $Mn_4Ca_1O_5$ cluster. Because ChlF seems likely to function as a simple homodimer and belongs to the earliest diverging clade of PsbA sequences in phylogenetic analyses, Chl *f* synthase may have been a transitional ancestor of water-oxidizing PSII. This hypothesis provides a simple explanation for the occurrence of multiple reaction centers in an ancestral cell. Thus, a Chl *a* photo-oxidoreductase that initially evolved for enhanced utilization of FRL may explain the origin of oxygen-evolving PSII. From an applied perspective, knowing the identity of ChlF may provide a tractable route for introducing the capacity for FRL utilization into crop plants, which would greatly expand the wavelength range for their solar energy utilization.



Evolutionary scheme for reaction centers (RC) in cyanobacteria. A primordial homodimeric RC (gray) led to type-1 (olive) and type-2 (cornflower) RCs by gene duplication—leading to photosystems I and II (PSI and PSII). Gene duplication/divergence events in the type-2 lineage (dotted arrows) possibly led to homodimeric chlorophyll *f* synthase (ChlF; blue) and a heterodimeric PSII ancestor (blue/magenta). Alternatively, the D1 (PsbA) subunit of PSII lost its $Mn_4Ca_1O_5$ cluster, leading to ChlF. (71 words)



Structural model of ChlF homodimer based on PsbA of PSII (PDB XXXX). The PsbA4/ChlF homodimer polypeptides are shown as blue ribbon structures. Putative substrate chlorophyll (Chl) *a* molecules are shown in bright green at the bottom. Cofactors are Chl *a* (dark green), pheophytin *a* (pale green), plastoquinones (burgundy), β-carotene (yellow) and tyrosine Y_Z (magenta). One Chl *a* molecule has been omitted for clarity. (63 words)

Introduction

Terrestrial cyanobacteria are an understudied and under-appreciated component of global primary productivity (GPP). About 55% of GPP is estimated to occur on land (*1*), and of the 45% of GPP that occurs in the oceans, oxygen-evolving *Prochlorococcus* and *Synechococcus* spp. are responsible for about ~50% of the total (*2, 3*). Given that the estimates for total marine and terrestrial cyanobacterial biomass are similar (*4*), terrestrial cyanobacteria could account for a substantial proportion of GPP and nitrogen fixation (*4, 5*). Terrestrial cyanobacteria often occur in environments that receive strongly filtered light because of shading by plants or because of their associations with soil crusts, benthic mat communities, or dense cyanobacterial blooms (*6*). The light available in such environments becomes highly enriched in far-red light (FRL), specifically wavelengths >700 nm (*6*). Therefore, cyanobacteria that are able to utilize FRL for photosynthesis gain a strong selective advantage over organisms that are unable to do so (*6-8*).

Cyanobacteria that grow in FRL undergo an extensive photoacclimation process, denoted far-red light photoacclimation (FaRLiP) (*7*). FaRLiP involves the FRL-dependent expression of a conserved cluster of twenty genes (Fig. S1). Seventeen of these genes are paralogs of genes encoding core subunits of photosynthetic complexes that are expressed in white light (WL). The expression of the paralogous genes in FRL leads to extensive remodeling of the cores of the major light-absorbing complexes of cyanobacteria: phycobilisomes (*9*), Photosystem (PS) I (*10*), and PSII (*6-8, 11*). About 10% of the chlorophyll (Chl) *a* molecules are replaced by Chl *f* in the paralogous PSI and PSII complexes produced in FRL are replaced by Chl *f* (*7, 8*), and allophycocyanin-B variants absorbing beyond 700 nm are also synthesized during FaRLiP (*7, 8, 12, 13*). Together these changes allow cells to utilize FRL for growth (*7, 8*). The FaRLiP gene cluster also contains three regulatory genes, *rfpA, rfpB,* and *rfpC*, which form a signaling cascade

that controls the expression of the FaRLiP genes (*14*). RfpA is a knotless RL/FRL-responsive phytochrome (RfpA) with a histidine kinase domain, RfpC is a small CheY-like response regulator that probably acts as a phosphate shuttle, and RfpB is a response regulator with a winged-helix, DNA-binding domain and two CheY-like receiver domains (*7, 14*). Null mutants for any one of these three *rfp* genes cannot grow in FRL, cannot express the genes in the FaRLiP gene cluster, and cannot synthesize Chl *f* (*14*).

Chl *f*, i.e., 2-formyl Chl *a*, was only recently discovered (*15–17*) and is synthesized when cells capable of FaRLiP are grown in FRL (*6-8, 18-20*). Chl *f* absorbs maximally at ~707 nm in 100% methanol (*21, 22*), and it is presumably made by reaction(s) extending from Chl *a* or chlorophyllide *a*, the immediate precursor of Chl *a* (Fig. 1) (*17, 23*). Chl *f* could presumably be introduced into crop plants to extend their light harvesting into the far-red (700 to 800 nm) and thereby improve their photosynthetic light-use efficiency (*24-26*). However, the enzyme(s) responsible for Chl *f* synthesis has not been identified, and the complexity of Chl *f* synthesis remains unknown.

Identification of the gene encoding Chl *f* synthase by reverse genetics

To identify the Chl *f* synthase, the enzyme responsible for the synthesis of Chl *f*, we employed reverse genetics in two cyanobacteria capable of FaRLiP, *Chlorogloepsis fritschii* PCC 9212 and *Synechococcus* sp. PCC 7335, and heterologous gene expression in the model cyanobacterium *Synechococcus* sp. PCC 7002 (hereafter *Synechococcus* 7002). Several observations led us to focus on *psbA4*, which encodes a so-called "super-rogue" paralog (sr-PsbA) of PsbA, the D1 core subunit of PSII (*27*). A sr-*psbA4* gene is found in each FaRLiP gene cluster (see Fig. S1) (*7, 8*), and *psbA4* expression is under the control of the RfpABC sensor

kinase/response regulator system (*14*). Transcription profiling of FaRLiP strains (e.g., *7, 14*) revealed that relative transcript abundances increase dramatically for hundreds of genes when cells are shifted from white light (WL) into FRL. However, phylogenetic profiling of ~600 genes whose relative transcript abundances increased more than twofold in FRL in *Leptolyngbya* sp. JSC-1 (*7*), and in particular monooxygenases, oxidoreductases, and other enzymes that might be logical candidates to be Chl *f* synthase, showed that only the genes of the FaRLiP cluster are universally present in cyanobacteria capable of performing FaRLiP. These observations strongly suggested that the gene(s) responsible for Chl *f* synthesis must be present in the FaRLiP gene cluster. Sr-PsbA4 is predicted to be structurally similar to the PsbA core subunit of PSII reaction centers, but Sr-PsbA4 sequences collectively lack key residues required for binding the $Mn_4Ca_1O_5$ cluster for water oxidation (*27, 28*) (Figs. 2C, S2). Additionally, sr-PsbA4 has some important differences in the vicinity of their putative plastoquinone binding site (*8, 27, 28*) (Fig. S2). However, sr-PsbA4 has retained tyrosine Y_Z and the ligands for binding Chl *a* (Figs. 2, S2). The structural and functional changes imposed by the inability to bind the $Mn_4Ca_1O_5$ cluster strongly suggested that sr-PsbA4 has a function other than being a core subunit of PSII.

Using a conjugation-based DNA transfer system (*14*), we constructed null mutants for *psbA4* genes in two cyanobacteria capable of FaRLiP (*8*), *Chl. fritschii* PCC 9212 and *Synechococcus* sp. PCC 7335 (Fig. S3). Neither *psbA4* mutant was able to synthesize Chl *f* when the mutant cells were grown in FRL. The characteristic, long-wavelength fluorescence emission at ~740-750 nm associated with PS complexes containing Chl *f* was completely missing for the two *psbA4* mutants (Fig. 3D and Fig. S4D). HPLC analyses of pigments extracted from cells of the *psbA4* null mutants grown in FRL confirmed the absence of Chl *f* and showed that only Chl *a* and Chl *d* were synthesized (Fig. 4 and Fig. S5). RT-PCR analyses showed that this was not due

to a loss of transcriptional control by RfpABC. Transcripts of selected FaRLiP genes exhibited normal RL/FRL-dependent regulation (Fig. 5). PsaA2 and PsaB2 were detected at low levels in PSI trimers isolated from *Chl. fritschii* PCC 9212 by proteomic analysis (Table S1). In the absence of Chl *f* in the *psbA4* mutant, it seems that FaRLiP subunits accumulate to much lower levels than in wild-type cells grown in FRL (Table S1).

Heterologous expression of the sr-*psbA4* gene in *Synechococcus* 7002.

To ascertain directly if sr-PsbA4 is the Chl *f* synthase, we expressed the *Chl. fritschii* PCC 9212 *psbA4* gene heterologously under the control of the strong *Synechocystis* sp. PCC 6803 *cpcBA* promoter from plasmid pAQ1 in the model cyanobacterium *Synechococcus* 7002 (*29*). *Synechococcus* 7002 does not naturally synthesize Chl *f* and is unable to grow in FRL (*7*). Cells were grown at low irradiance in WL (15 µmol photons m^{-2} s^{-1}) at room temperature in hopes of minimizing sr-PsbA4 turnover and preventing photooxidation of any Chl *f* that might be formed, because no known Chl *f*-binding proteins are present in this heterologous system. Reversed-phase HPLC analysis of pigments extracted from wild-type *Synechococcus* 7002 cells showed only Chl *a* as expected (Fig. 6A). However, HPLC analysis showed that pigments extracted from the *Synechococcus* 7002 strain expressing *psbA4* contained both Chl *a* and Chl *f* (Fig. 6A, Fig. S6). Although only a small amount of Chl *f* was produced (~0 .059% of total Chl), this pigment had the same retention time as Chl *f* from cells of *Chl. fritschii* PCC 9212 grown in FRL (Fig. S6). Moreover, the in-line absorption spectrum of the pigment from the *Synechococcus* 7002 strain expressing *psbA4* was identical to those of Chl *f* from three FaRLiP strains (Fig. 6B). Finally, MS analysis of the new pigment, after treatment with formic acid to produce the corresponding pheophytin, confirmed that this pigment had the expected *m/z* values

of 885.4 $[M + H]^+$, 902.4 $[M + NH_4^+]^+$, and 907.4 $[M + H]^+$ for Chl *f* (Fig. S7). Thus, expression of the *psbA4* gene is sufficient to direct the synthesis of Chl *f* in *Synechococcus* 7002. Collectively, the results show that the *psbA4* gene encodes Chl *f* synthase, and accordingly, we propose to rename this gene *chlF*.

Light dependence of Chl *f* synthase reaction

The discovery that a paralog of the D1 core subunit of PSII was responsible for Chl *f* synthesis was unexpected, because the implication was that this enzyme must be structurally and functionally related to PSII. Interestingly, the oxidation of Chl *a* to Chl *f* (or Chlide *a* to Chlide *f*) is a 4-electron oxidation—like water oxidation to produce dioxygen (*30*). We propose that ChlF is a photo-oxidoreductase that uses light to oxidize Chl *a* (or Chlide *a*) to produce Chl *f* (or Chlide *f*) while reducing bound plastoquinone (Figs. 1, 2). If this hypothesis is correct, Chl *f* synthesis should be minimal if cells are allowed to express *chlF* but then shifted into darkness. To test this hypothesis, we grew wild-type *Synechococcus* 7002 cells expressing *chlF* in a medium containing 10 mM glycerol, a metabolizable carbon/energy source that supports heterotrophic growth (*31*), in continuous FRL, continuous darkness, and in low-intensity WL (for details, see **Materials and Methods**). There was a low level of Chl *f* in the cells at time zero (Fig. 7), and this level decreased slightly after 3 days in darkness, although the cells grew and their Chl *a* content increased. Similar amounts of Chl *a* were synthesized by cells grown for 3 days in continuous low-intensity WL or continuous FRL, and both synthesized identical amounts of Chl *f* (Fig. 7). The absorption spectrum of the pigment eluting at 14.5 min in Fig. 7 was identical to the spectra shown in Fig. 6B (Note that retention times are different because a

different HPLC protocol was employed compared to Fig. 6A; see Materials and Methods for details).

We also performed an alternative experiment designed to test whether light is required for the synthesis of Chl *f* (Fig. 8). Wild-type *Chl. fritschii* PCC 9212 was grown in continuous FRL or FRL interrupted by a dark period of 12 h. At time zero, *Chl. fritschii* PCC 9212 WT cells that had been grown in a medium containing 5 mM fructose, a metabolizable carbon/energy source, were inoculated into two cultures, which were transferred to FRL for 24 to induce the expression of the FaRLiP gene cluster. One of the cultures was left in continuous FRL (Fig. 8A), while the other placed in darkness for 12 h and then returned to FRL for another 12 h (Fig. 8B). The Chl *f* content of the control cells, which could be detected by the increasing 740-nm low-temperature fluorescence emission from Chl *f* increased throughout the FRL incubation. In the other sample, Chl *f* synthesis stopped when cells were placed in darkness, but Chl *f* synthesis resumed when cells were returned to FRL (Fig. 8B).

The data from both of these experimental approaches demonstrate that light is required for the synthesis of Chl *f* by ChlF/sr-PsbA4. Thus, we conclude that ChlF is a photo-oxidoreductase. Chl *f* synthase is the second light-dependent enzyme of Chl biosynthesis, the other being light-dependent protochlorophyllide oxidoreductase (POR) (*32*).

Implications of Chl *f* synthase (ChlF) on evolution of water oxidation and PSII

Within the context of PSII evolution, there are two ways to view the result that ChlF/PsbA4 is an enzyme that probably photooxidizes Chl *a* (or Chlide *a*) to produce Chl *f* (or Chlide *f*). In the first scenario, ChlF may have evolved from the PsbA (D1) subunit of a water-oxidizing PSII complex by loss of the ligands for binding the $Mn_4Ca_1O_5$ cluster but retention of

catalytically useful chlorophylls, tyrosine Y_Z, and plastoquinone binding after gene duplication and divergence. Although this is certainly a reasonable possibility, phylogenetic analyses do not support this hypothesis (*27, 28*). An alternate possibility is perhaps more appealing: ChlF/PsbA4 could be ancestral to PsbA of PSII, and PsbA could have arisen by gene duplication and divergence to bind a $Mn_4Ca_1O_5$ cluster to catalyze water oxidation, thus providing an explanation for the origin of PSII (Fig. S8) (*28*). Considering that the selection pressure to expand the wavelength range of photosynthesis is largely responsible for the many different Chls and antenna complexes extant today, this scheme provides an explanation for the origin of ChlF, for the reaction that it catalyzes, and for the occurrence of a primitive type-2 reaction center in an ancestral cyanobacterium that may also have had a PSI-like reaction center (*33, 34*).

Supporting this second possibility further, phylogenetic analyses show that ChlF/sr-PsbA, together with a highly unusual PsbA-like sequence from *Gloeobacter kilaueensis*, are the earliest diverging members of the PsbA family (Fig. S8) (*8, 27, 28*). We hypothesize that ChlF functions as homodimer, because introduction of *psbA4/chlF* into *Synechococcus* 7002 was sufficient to enable Chl *f* synthesis. All currently known photochemical reaction centers are either homodimers or heterodimers (*33, 34*). Furthermore, because ChlF is apparently a photo-oxidoreductase and oxidation of the substrate probably occurs by a radical mechanism, assuming a one-photon, one-electron mechanism as in PSII (*30*), dioxygen may not be required for the synthesis of Chl *f*. Thus, it is possible that Chl *f* synthase appeared before the atmosphere became oxic (i.e., before PSII had evolved). These implications could mean that ChlF is a long-sought, transitional intermediate in the evolution of oxygenic photosynthesis: a simple, homodimeric, type-2 reaction center that might have first evolved in anoxygenic ancestors of modern cyanobacteria (Fig. S8) (*28, 34, 35*). Subsequent gene duplication events (two are required to

form the heterodimers found in PSII; *35*) and divergence to acquire the ligands to bind the $Mn_4Ca_1O_5$ cluster, could have ultimately led to an enzyme that could oxidize Chl (or Chlide) to an enzyme that could oxidize water.

In either case, the likely selective pressure for the synthesis of Chl *f* would have been to extend photosynthetic light harvesting into the far-red spectral region. The ability to use FRL or near-infrared light for photosynthesis provides a strong selective advantage to those organisms that can extend their light-harvesting range beyond the wavelengths absorbed by Chl *a* (*7*). Examples include the homodimeric type-1 reaction centers of *Chloracidobacterium thermophilum* (*36*) and green sulfur bacteria (*37*), both of which bind Chl *a* and bacteriochlorophyll (BChl) *a*. Similarly, the unusual cyanobacterium, *Acaryochloris marina*, has evolved to use Chl *d*, another far-red absorbing Chl (see Fig. 1), together with Chl *a* to extend its photosynthetic light harvesting into the far-red region (*38, 39*). It is increasingly clear that Chlide *a* is the central, "hub" compound of Chl biosynthesis, and that the biosynthetic pathway repeatedly diverged away from this molecule, or one of its precursors 3,8-divinyl-protochlorophyllide, to extend light harvesting into the far-red/near infra-red (*23, 40*). These findings strongly support the Granick hypothesis concerning the evolution of heme and Chl biosynthesis, namely that pathways evolved forward as organisms evolved (see discussion in *40*).

By using domain-swapping and site-directed mutagenesis methods, it may be possible to recapitulate evolution to determine which amino acid residues must be changed in PsbA to obtain a protein that no longer binds a $Mn_4Ca_1O_5$ cluster but can synthesize Chl *f* (or more interestingly, the converse). Such approaches could provide an experimental framework for studying the origins of PSII, water oxidation, and thus oxygenic photosynthesis. Finally, because Chl *f* biosynthesis is the product of a single gene product, it may now be feasible to introduce the

capacity for Chl *f* biosynthesis into plants to extend the wavelength range for their photosynthesis into the far-red region of the solar spectrum (*24-26*).

Materials and Methods

Organisms and cultivation conditions

Cyanobacterial strains used in this study were obtained from the Pasteur Culture Collection (http://www.pasteur.fr/pcc_cyanobacteria) (*41*). *Synechococcus* 7002 was grown in medium A containing 10 mM nitrate (denoted as Medium A^+) (*42*). This cyanobacterium was routinely grown under "standard conditions" (see *43*): 38°C, sparging with 1% CO_2 (v/v) in air, at 250 µmol photons m^{-2} s^{-1} provided by cool white fluorescent tubes. For some experiments, cells were grown at room temperature at an irradiance of ~15 µmol photons m^{-2} s^{-1}.

Synechococcus sp. PCC 7335 was cultured in ASN-III growth medium (*41*) to which vitamin B_{12} (10 µg ml^{-1}, final concentration) and Tris-HCl, pH 8.0 (10 mM, final concentration) were added. *Chlorogloeopsis fritschii* PCC 9212 was grown in the B-HEPES growth medium (*44*), a modified BG-11 medium containing 1.1 g L^{-1} HEPES (final concentration) with the pH adjusted to 8.0 with 2.0 M KOH. Cool white fluorescent bulbs provided continuous illumination at ~250 µmol photons m^{-2} s^{-1} (WL), and liquid cultures were sparged with 1% (v/v) CO_2 in air. *Synechococcus* sp. PCC 7335 was grown at a lower WL intensity of ~50 µmol photons m^{-2} s^{-1}. In some experiments, *Chl. fritschii* PCC 9212 was grown mixotrophically by adding 5 mM fructose to the growth media. Far-red light was provided by Epitex, L720-06AU LEDs (Marubeni, Santa Clara, CA) with emission centered at 720 nm (26-30 µmol photons m^{-2} s^{-1}), in combination with plastic filters as previously described (*7, 14*). Light boards composed of 50 LEDs were assembled. Although all should nominally emit "720 nm" light, the peak emission ranges from 705 nm to 735 nm. Furthermore, the bandwidth characteristics for these LEDs is such that light from 660 nm to 760 nm is actually produced (see www.epitex.com/products/lpm/pdfs/L720-06AU.pdf). Growth of cells was monitored

turbidometrically at 750 nm by using a GENESYS 10 spectrophotometer (ThermoSpectronic, Rochester, NY).

Media for the growth of mutants of *Chl. fritschii* PCC 9212 and *Synechococcus* sp. PCC 7335 and the *psbA4* overexpression strain of *Synechococcus* 7002 were amended with antibiotics as required: kanamycin (100 μg ml^{-1}); erythromycin (20 μg ml^{-1}); gentamicin (50 μg ml^{-1}).

Cultivation conditions to demonstrate the light requirement for Chl *f* synthase

To assess the requirement for light for Chl *f* biosynthesis in a strain of *Synechococcus* 7002 expressing *chlF/psbA4* from *Chl. fritschii*, a starter culture was grown in high light for 24 hours in medium A^+ containing 10 mM glycerol. This starter culture was used to inoculate duplicate cultures that were grown in the same medium under continuous FRL (28 μM photons m^{-2} s^{-1}), low WL (15 μM photons m^{-2} s^{-1}), and continuous darkness. The data shown in Fig. 7 indicate that only cells maintained in FRL or low-intensity WL could synthesize Chl *f*. However, cells grew under all three conditions. The $OD_{730\ nm}$ of the culture grown in darkness increased ~30%, but the $OD_{730\ nm}$ of cells grown in continuous FRL increased ~90% and that for cells grown in continuous WL ~60%. Importantly, the Chl *a* content of cells grown in darkness also increased during incubation in the dark, showing that these cells were capable of synthesizing Chl *a* but not Chl *f* (Fig. 7).

In a second experiment to test the requirement for light for Chl *f* biosynthesis, the WT strain of *Chl. fritschii* PCC 9212 was grown in continuous or interrupted FRL (Fig. 8). At time zero WT cells grown in WL in a medium containing 5 mM fructose, which supports heterotrophic growth of *Chl. fritschii* PCC 9212 (*41*), were inoculated into two cultures, which were transferred to FRL for 24 h. One of the cultures was left in continuous FRL (Fig. 8A),

while the other was placed in darkness for 12 h and then returned to FRL for another 12 h (Fig. 8B). The control cells incubated in continuous far-red light continuously synthesized Chl *f,* which could be detected by the increasing 740-nm fluorescence emission maximum from Chl *f*. Over the 48 duration of the experiment, the cell density under each growth condition increased by about 25%. Chl *f* synthesis stopped when cells were placed in darkness, but Chl *f* synthesis resumed when cells were returned to FRL (Fig. 8B). This experiment was repeated twice with essentially identical results for both trials. This experiment confirms that Chl *f* synthase requires light for its activity.

Construction of mutants and conjugation

Professor Jindong Zhao from Peking University provided the biparental conjugation system (*45*), and the method employed was very similar to that described in (*14*). The donor *E. coli* strain HB101 contained the conjugal plasmid pRL443 and the helper plasmid pRL623 (*46, 47*). Supplemental Fig. S3 shows maps of the scheme employed to delete the *psbA4* genes of *Chl. fritschii* PCC 9212 and *Synechococcus* sp. PCC 7335. The *psbA4* gene of *Chl. fritschii* was deleted and replaced by an *ermC* cassette conferring resistance to erythromycin, while the *psbA4* gene of *Synechococcus* sp. PCC 7335 was partly deleted and replaced by an *aphII* cassette, conferring resistance to kanamycin. Upstream and downstream DNA fragments for each target *psbA4* gene were cloned into the cargo plasmid pRL277 with an interposing DNA fragment encoding *ermC* or *aphII* to replace the target gene through a double-crossover recombination event. For each target gene, upstream and downstream DNA fragments with a size of ~2.5 to 3.0 kb were amplified by PCR using Phusion® HF DNA polymerase (New England Biolabs,

Ipswich, MA, USA). The primers for DNA fragment amplification of the upstream and downstream regions are listed in Table S2.

Cargo plasmids were transformed into the donor *E. coli* HB101 cells containing the conjugal plasmid pRL443 and the helper plasmid pRL623. The resulting *E. coli* HB101 strains were grown in 5 to 20 ml of Luria-Bertani (LB) medium supplemented with appropriate antibiotics and cultured at 37°C overnight. The *E. coli* cells were harvested by centrifugation at low speed, washed with fresh LB medium 3 times, and resuspended in fresh LB, B-HEPES medium, or ASN-III medium (100-200 μl). Meanwhile, freshly grown wild-type *Chl. fritschii* 9212 or *Synechococcus* sp. PCC 7335 cells ($OD_{750\ nm}$ = 0.6-1.0; 500 μl to 2.0 ml) were centrifuged at low speed, washed with fresh B-HEPES or ASN-III medium 3 times, and finally resuspended in B-HEPES or ASN-III medium (100 to 200 μl). The *E. coli* and cyanobacterial cells were gently mixed in a sterile microcentrifuge tube and incubated at 30°C under low light for 4-6 hours, and the cell mixture was then spread onto a sterile nitrocellulose filter overlaid on a B-HEPES or ASN-III agar plate without antibiotics. The plate was incubated at 30°C under low light for another 18-36 hours, and then the filter was transferred to a B-HEPES agar plate containing 20 μg erythromycin ml^{-1} or an ASN-III agar plate containing 100 μg kanamycin ml^{-1}. Brown- or green-colored colonies appeared on the filter after about four weeks. Colonies were picked and streaked repeatedly on selective medium, and the purified transconjugant cells were ultimately transferred to liquid B-HEPES or ASN-III medium for analyses. Segregation of wild-type and mutant alleles of the target gene was analyzed by colony PCR.

Pigment extraction, HPLC analyses, and mass spectrometry

Two methods, denoted MY1 and MY2, were used for pigment extraction and HPLC analysis. Method MY1 was performed as previously described (*7, 8, 14*). Pigments were extracted with acetone/methanol (7:2, v/v) from cells and analyzed by reversed-phase HPLC on an Agilent 1100 HPLC system (Agilent Technologies, Santa Clara, CA) with an analytical Discovery C18 column (4.6 mm × 25 cm) (Supelco, Sigma-Aldrich, St. Louis, MO). The gradient elution program [B, minutes] using Solvent A (methanol: acetonitrile: H_2O = 42: 33: 25) and Solvent B (methanol: acetonitrile: ethyl acetate = 50: 20: 30) was set as [30%, 0], [100%, 50], [100%, 58], [30%, 60] at a flow rate 1 mL min^{-1}. Elution of pigments was monitored at 705 nm, the absorbance maximum of Chl *f*.

Method MY2, adapted from Armenta *et al*. (*48*), was performed for both analytical and semi-preparative purposes. Pigments were extracted using a previously published method (*49*) and separated with the same HPLC system using an analytical Discovery C18 column described above or a semi-preparative Discovery C18 column (10 mm × 25 cm) (Supelco, Sigma-Aldrich, St. Louis, MO). Solvents A and B were 88:10:2 methanol/ethyl acetate/H_2O (v/v/v) and 48:50:2 methanol/ethyl acetate/H_2O (v/v/v), respectively. After 10 min of washing with Solvent A, Chls were eluted with a linear gradient of 0% to 20% solvent B over 20 min, increasing to 100% to wash the column. The analytical program was performed at 1 mL min^{-1} and 3.5 mL min^{-1} for analytical and semi-preparative purposes, respectively.

Mass spectrometric analysis of chlorophyll pigments collected by semi-preparative method MY2 was performed on a Waters Q-TOF Premier quadrupole/time-of-flight (TOF) mass spectrometer (Waters Corporation (Micromass Ltd.), Manchester, UK). Operation of the mass spectrometer was performed using MassLynx™ software Version 4.1 (http://www.waters.com). Samples were introduced into the mass spectrometer using a Waters 2695 HPLC system. The

samples were analyzed using flow injection analysis (FIA), in which the sample is injected into the mobile phase flow and passed directly into the mass spectrometer, where the analytes are ionized and detected. The mobile phase used was 90% acetonitrile (LC-MS grade) and 10% aqueous 0.1% formic acid. The flow rate was 0.15 mL min^{-1}. The nitrogen drying gas temperature was set to 300 °C at a flow of 6 L min^{-1}, and the capillary voltage was 2.8 kV. The mass spectrometer was set to scan from 100-1000 *m/z* in positive ion mode, using electrospray ionization (ESI). The MS analysis of Chls and pheophytins was performed in the Proteomics and Mass Spectrometry Core Facility, the Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park.

Absorption and fluorescence spectroscopy

Room-temperature absorption spectra were recorded with a GENESYS 10 spectrophotometer (ThermoFisher Scientific, Waltham, MA) or a Cary 14 UV-Vis-NIR spectrophotometer modified for computer-controlled operation by OLIS, Inc. (Bogart, GA). Absorption spectra of whole cells of *Synechococcus* sp. PCC 7335 and mutants derived from it were recorded in growth medium ASN-III. The absorption spectra of *Chl. fritschii* PCC 9212 and mutant strains derived from it were recorded by homogenizing, resuspending and diluting cells in B-HEPES medium with 60% w/v sucrose (one-to-ten dilution). Prior to recording the absorption spectra, whole cells were homogenized as necessary, and spectra were recorded through the opal glass side of the cuvettes to allow correction for scattering. Fluorescence emission spectra of whole cells at 77 K were measured with an SLM 8000C spectrofluorometer, modified for computer-controlled operation by OLIS, Inc. (Bogart, GA). Cells in exponential growth phase ($OD_{750\ nm}$ = ~0.6 to 0.7) were collected, homogenized and resuspended in 50 mM Tris–HCl, pH

8.0 buffer. Glycerol was added to a final concentration of 60% (v/v). Cells were adjusted to a concentration of ~0.5 $OD_{750\ nm}$ and quickly frozen in liquid nitrogen. The excitation wavelength was set to 440 nm to excite Chls preferentially.

RNA isolation and RT-PCR

For gene expression analysis, total RNA was isolated the wild type and mutant strains grown to mid-exponential growth phase using a High Pure RNA isolation kit (Roche Diagnostics, Indianapolis, IN). To eliminate trace amounts of contaminating chromosomal DNAs, RNA samples were incubated with RNase-free DNase I (Promega, Madison, WI) for 2-3 h at 37°C. DNase-treated RNA samples were further purified using RNA purification cartridges. The absence of DNA in RNA samples was verified by PCR assays in which reverse transcriptase was omitted. The concentration of RNA was determined by NanoDrop measurements (Thermo Scientific, Watham, MA).

Reverse transcription-polymerase chain reaction (RT-PCR) analysis of total RNA isolated from wild-type and mutant strains was performed using a MyTaq One-Step RT-PCR kit (Bioline USA Inc., Taunton, MA). Primers were designed to amplify ~500-bp regions of specifically selected genes encoding major subunits of PSI, PSII and phycobilisomes, as well as the 16S RNA gene as a control (see legend to Fig. 5). The resulting amplicons were analyzed by agarose gel electrophoresis.

Expression of the *psbA4* gene of *Chl. fritschii* PCC 9212 in *Synechococcus* 7002

To generate the expression vector for *psbA4*, a DNA fragment containing the *psbA4* gene was amplified by PCR from chromosomal DNA of *Chl. fritschii* PCC 9212 using Phusion® HF

DNA polymerase. The primers used were 9212 *psbA4* pAQ-1-1 (NcoINdeI) and 9212 *psbA4* pAQ-1-2 (BamHI) (Table S2). After digestion with NdeI and BamHI, the *psbA4* gene amplicon was cloned into the pAQ1Ex-P_{cpcBA}/Gm shuttle vector (*29*). After verification by DNA sequencing, the *psbA4* expression construct (pAQ1Ex*::*9212*_psbA4* (Gm^R)) was transformed into the cells of *Synechococcus* 7002. Transformants were selected on agar plates made with A^+ medium amended with gentamicin. Transformants harboring the expression plasmid were confirmed by PCR analysis and DNA sequence analysis.

Isolation of PSI complexes and LC-MS-MS analysis

PSI complexes of the wild-type and *psbA4* mutant strains of *Chl. fritschii* PCC 9212 were isolated from cells grown under WL and FRL conditions following the procedures described previously (*7*). In-solution digestion of proteins with trypsin and LC-MS-MS analysis were performed at the PARC Mass Spectrometer Facility at Washington University in St. Louis. Peptide mass spectra (*m/z* range 380-1500) were acquired at high mass resolving power (70,000 for ions with *m/z* = 200) with a Fourier Transform (FT) Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The raw data from the LC-MS-MS analysis was directly loaded into PEAKS (v 7.0, Bioinformatics Solution Inc., Waterloo, ON, Canada) for performing database searches against the total proteome for *Chl. fritschii* PCC 9212, which was derived from the annotated genome sequence.

Supplementary Materials
www.sciencemag.org
Figures S1 – S8
Tables S1 and S2

References and Notes

1. P. G. Falkowski, J. A. Raven *Aquatic Photosynthesis,* Princeton University Press, Princeton, NJ (2013)
2. M. R. Speight, P. A. Henderson, *Marine Ecology: Concepts and Applications*, Wiley-Blackwell, West Sussex, UK (2010)
3. P. Flombaum, J. L. Gallegos, R. A. Gordillo, J. Rincón, L. L. Zabala, N. Jiao, D. M. Karl, W. K. W. Li, M. W. Lomas, D. Veneziano, C. S. Vera, J. A. Vrugt, A. C. Martiny, Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. USA* **110**, 9824–9829 (2013).
4. F. Garcia-Pichel, J. Belnap, S. Neuer, F. Schanz, Estimates of global cyanobacterial biomass and its distribution. *Arch. Hydrobiol.* Suppl. **148**, 213–227 (2003).
5. W. Elbert. B. Weber, S. Burrows, J. Steinkamp, B. Büdel, M. O. Andreae, and U. Pöschl, Contribution of cryptogamic covers to the global cycles of carbon and nitrogen. *Nature Geosci.* **5**, 459–462 (2012).
6. F. Gan, D. A. Bryant, Adaptive and acclimative responses of cyanobacteria to far-red light. *Environ. Microbiol.* **17**, 3450–3465 (2015).
7. F. Gan, S. Zhang, N. C. Rockwell, S. S, Martin, J. C. Lagarias, D. A. Bryant, Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light. *Science* **345**, 1312–1317 (2014).

8. F. Gan, G. Shen, D. A. Bryant, Occurrence of far-red light photoacclimation (FaRLiP) in diverse cyanobacteria. *Life* **5**, 4–24 (2015).

9. W. A. Sidler, Phycobilisome and phycobiliprotein structures, in *The Molecular Biology of Cyanobacteria*, D. A. Bryant, Ed., *Advances in Photosynthesis and Respiration*, Vol. 1 (Kluwer Academic, Dordrecht, The Netherlands, 1994), pp. 139–216.

10. I. Grotjohann, P. Fromme, Structure of cyanobacterial photosystem I. *Photosynth. Res.* **85**, 51–72 (2005).

11. Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55–60 (2011).

12. A. N. Glazer, D. A. Bryant, Allophycocyanin B (λ_{max} 671, 618 nm): a new cyanobacterial phycobiliprotein. *Arch. Microbiol.* **104**, 15-22 (1975).

13. Y. Li, Y. Lin, C. J. Garvey, D. Birch, R. W. Corkery, P. C. Loughlin, H. Scheer, R. D. Willows, M. Chen, Characterization of red-shifted phycobilisomes isolated from the chlorophyll *f*-containing cyanobacterium *Halomicronema hongdechloris*. *Biochim. Biophys. Acta* **1857**, 107–114 (2016).

14. C. Zhao, F. Gan, G. Shen, D. A. Bryant, RfpA, RfpB, and RfpC are the master control elements of far-red light photoacclimation (FaRLiP). *Front. Microbiol.* **6**, 1303 (2015).

15. M. Chen, M. Schliep, R. D. Willows, Z.-L. Cai, B. A. Neilan, H. Scheer, A red-shifted chlorophyll. *Science* **329**, 1318–1319 (2010).

16. R. D. Willows, Y. Li, H. Scheer, M. Chen, Structure of chlorophyll *f*. *Org. Lett.* **15**, 1588–1590 (2013).

17. M. Chen, Chlorophyll modifications and their spectral extension in oxygenic photosynthesis. *Annu. Rev. Biochem.* **83**, 317–340 (2014).

18. Y. Li, Y. Lin, P. C. Loughlin, M. Chen, Optimization and effects of different culture conditions on growth of *Halomicronema hongdechloris*—a filamentous cyanobacterium containing chlorophyll *f*. *Front. Plant Sci.* **5**, 67 (2014).

19. R. L. Airs, B. Temperton, C. Sambles, G. Farnham, S. C. Skill, C. A. Llewellyn, Chlorophyll *f* and chlorophyll *d* are produced in the cyanobacterium *Chlorogloeopsis fritschii* when cultured under natural light and near-infrared radiation. *FEBS Lett.* **588**, 3770–3777 (2014).

20. L. Behrendt, A. Brenjnrod, M. Schliep, S. J. Sørensen, A. W. Larkum, M. Kühl, Chlorophyll *f*-driven photosynthesis in a cavernous cyanobacterium. *ISME J.* **9**, 2108–2011 (2015).

21. Y. Li, Z.-L. Cai, M. Chen, Spectroscopic properties of chlorophyll *f*. *J. Phys. Chem. B* **117**, 11309–11317 (2013).

22. Y. Li, N. Scales, R. E. Blankenship, R. D. Willows, M. Chen, Extinction coefficient for red-shifted chlorophylls: chlorophyll *d* and chlorophyll *f*. *Biochim. Biophys. Acta* **1817**, 1292–1298 (2012).

23. A. Gomez Maqueo Chew, D. A. Bryant, Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity. *Annu. Rev. Microbiol.* **61**, 113–129 (2007).

24. M. Chen, R. E. Blankenship. Expanding the solar spectrum used by photosynthesis. *Trends Plant Sci.* **16**, 427–431 (2011).

25. R. E. Blankenship, M. Chen, Spectral expansion and antenna reduction can enhance photosynthesis for energy production. *Curr. Opin Chem. Biol.* **17**, 457–461 (2013).

26. R. E. Blankenship, D. M. Tiede, J. Barber, G. W. Brudwig, G. Fleming, M. Ghirardi, M. R. Gunner, W. Junge, D. M. Kramer, A. Melis, T. A. Moore, C. C. Moser, D. G. Nocera, A. J.

Nozik, D. R. Ort, W. W. Parson, R. C. Prince, R. T. Sayer, Comparing photosynthetic and photovoltaic efficiences and recognizing the potential for improvement. *Science* **332**, 805–809 (2011).

27. J. W. Murray, Sequence variation at the oxygen evolving centre of photosystem II: a new class of 'rogue' cyanobacterial D1 proteins. *Photosynth. Res.* **110**, 177–184 (2012).

28. T. Cardona, J. W. Murray, A. W. Rutherford, Origin and evolution of water oxidation before the last common ancestor of the cyanobacteria. *Mol. Biol. Evol.* **32**, 1310–1328 (2015).

29. Y. Xu, R. M. Alvey, P. O. Byrne, J. E. Graham, G. Shen, D. A. Bryant, Expression of genes in cyanobacteria: adaptation of endogenous plasmids as platforms for high-level gene expression in *Synechococcus* sp. PCC 7002. *Methods Mol. Biol.* **684**, 273–293 (2011).

30. J.-R. Shen, The structure of photosystem II and the mechanism of water oxidation in photosynthesis. *Annu. Rev. Plant Biol.* **66**, 23–48 (2015).

31. D. H. Lambert, S. E. Stevens, Jr., Photoheterotrophic growth of *Agmenellum quadruplicatum* PR-6. *J. Bacteriol.* **165**, 654–656 (1986).

32. M. Gabruk, B. Mysilwa-Kurdziel, Light-dependent protochlorophyllide oxidoreductase: phylogeny, regulation and catalytic properties. *Biochemistry* **54**, 5255–5262 (2015).

33. M. F. Hohmann-Marriott, R. E. Blankenship, Evolution of photosynthesis. *Annu. Rev. Plant Biol.* **62**, 515–548 (2011).

34. T. Cardona, A fresh look at the evolution and diversification of photochemical reaction centers. *Photosynth. Res.* **126**, 111–134 (2015).

35. T. Cardona, Reconstructing the origin of oxygenic photosynthesis: do assembly and photoactivation recapitulate evolution? *Front. Plant Sci.* **7**, 257 (2016).

36. Y. Tsukatani, S. P. Romberger, J. H. Golbeck, D. A. Bryant, Isolation and characterization of homodimeric type-I reaction center complex from *Candidatus* Chloracidobacterium thermophilum, an aerobic chlorophototroph. *J. Biol. Chem.* **287**, 5720–5732 (2012).

37. G. Hauska, T. Schoedl, H. Remigy, G. Tsiotis, The reaction center of green sulfur bacteria. *Biochim. Biophys. Acta* **1507**, 260–277 (2001).

38. H. Miyashita, S. Ohkubo, H. Komatsu, Y. Sorimachi, D. Fukayama, D. Fujinuma, S. Akutsu, M. Kabayashi, Discovery of chlorophyll *d* in *Acaryochloris marina* and chlorophyll *f* in a unicellular cyanobacterium, strain KC1, isolated from Lake Biwa. *J. Phys. Chem. Biophys.* **4**, 149 (2014).

39. S. I. Allakhverdiev, V. D. Kreslavski, S. K. Zharmukhamedov, R. A. Voloshin, D. V. Korol'kova, T. Tomo, and J.-R. Shen, Chlorophylls *d* and *f* and their role in primary photosynthetic processes of cyanobacteria. *Biochemistry (Moscow)* **81**, 201–212 (2016).

40. D. A. Bryant, Z. Liu, Green Bacteria: insights into green bacterial evolution through genomic analyses. *Adv. Bot. Res.* **66**, 99–150 (2013).

41. R. Rippka, J. Deruelles, J. B. Waterbury, M. Herdman, R. Y. Stanier, Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* **111**, 1–61 (1979).

42. S. E. Stevens Jr., R. D. Porter, Transformation in *Agmenellum quadruplicatum*. *Proc. Natl. Acad. Sci. USA* **77,** 6052–6056 (1980).

43. M. Ludwig, D. A. Bryant, Transcription profiling of the model cyanobacterium *Synechococcus* sp. strain PCC 7002 by next-gen (SOLiD^{TM}) sequencing of cDNA. *Front. Microbiol.* **2**, 41 (2011) doi: 10.3389/fmicb.2011.00041.

44. J. M. Dubbs, D. A. Bryant, Molecular cloning and transcriptional analysis of the *cpeBA* operon of the cyanobacterium *Pseudanabaena* species PCC 7409. *Mol. Microbiol.* **5**, 3073–3085 (1991).

45. Y. Zhao, Y. Shi, W. Zhao, X. Huang, D. Wang, N. Brown, J. Brand, J. Zhao, CcpB, a calcium-binding protein from *Anabaena* sp. PCC 7120, provides evidence that calcium ions regulate heterocyst differentiation. *Proc. Natl. Acad. Sci. USA* **102**, 5744–5748 (2005).

46. J. Elhai, C. P. Wolk, Conjugal transfer of DNA to cyanobacteria. *Meth. Enzymol.* **167**, 747–754 (1988).

47. J. Elhai, A. Vepritskiy, A. M. Muro-Pastor, E. Flores, C. P. Wolk, Reduction of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain PCC 7120 *J. Bacteriol.* **179**, 1998–2005 (1997).

48. R. E. Armenta, A. Burja, H. Radianingtyas H, C. J. Barrow, Critical assessment of various techniques for the extraction of carotenoids and co-enzyme Q_{10} from the *Thraustochytrid* strain ONC-T18. *J. Agric. Food Chem.* **54**, 9752-9758 (2006).

49. D. P. Canniffe, P. J. Jackson, S. Hollingshead, M. J. Dickman, C. N. Hunter, Identification of an 8-vinyl reductase involved in bacteriochlorophyll biosynthesis in *Rhodobacter spaeroides* and evidence for the existence of a third distinct class of the enzyme. *Biochem. J.* **450**, 397-405 (2013).

50. M. Schliep, B. Crossett, R. D. Willows, and M. Chen, ^{18}O labeling of chlorophyll *d* in *Acaryochloris marina* reveals that chlorophyll *a* and molecular oxygen are precursors. *J. Biol. Chem.* **285**, 28450–28456 (2010).

51. 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).

52. R. L. Dunbrack, Jr., Rotamer libraries in the 21st century. *Curr. Opin. Struct. Biol*. **12**, 431–440 (2002).

Acknowledgments: This research was supported by National Science Foundation grant MCB-1021725 to D.A.B This research was also conducted under the auspices of the Photosynthetic Antenna Research Center (PARC), an Energy Frontier Research Center funded by the DOE, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC 0001035. M.-Y. H., G. S, C. Z., and D. A. B. were partly supported by PARC and partly supported by N. S. F. D. P. C. acknowledges support from a European Commission Marie Skłodowska-Curie Global Fellowship (660652). The authors would like to thank Drs. R. E. Blankenship, A. N. Glazer, J. H. Golbeck, C. N. Hunter, J. C. Lagarias, and J. T. J. Lecomte for reading, commenting, and offering suggestions to improve the draft manuscript. The authors thank J. Miller for his expert assistance in performing and interpreting the mass spectral data for Chls and pheophytins. The authors thank H. Zhang for assistance with the proteomic analyses.

Figure Legends

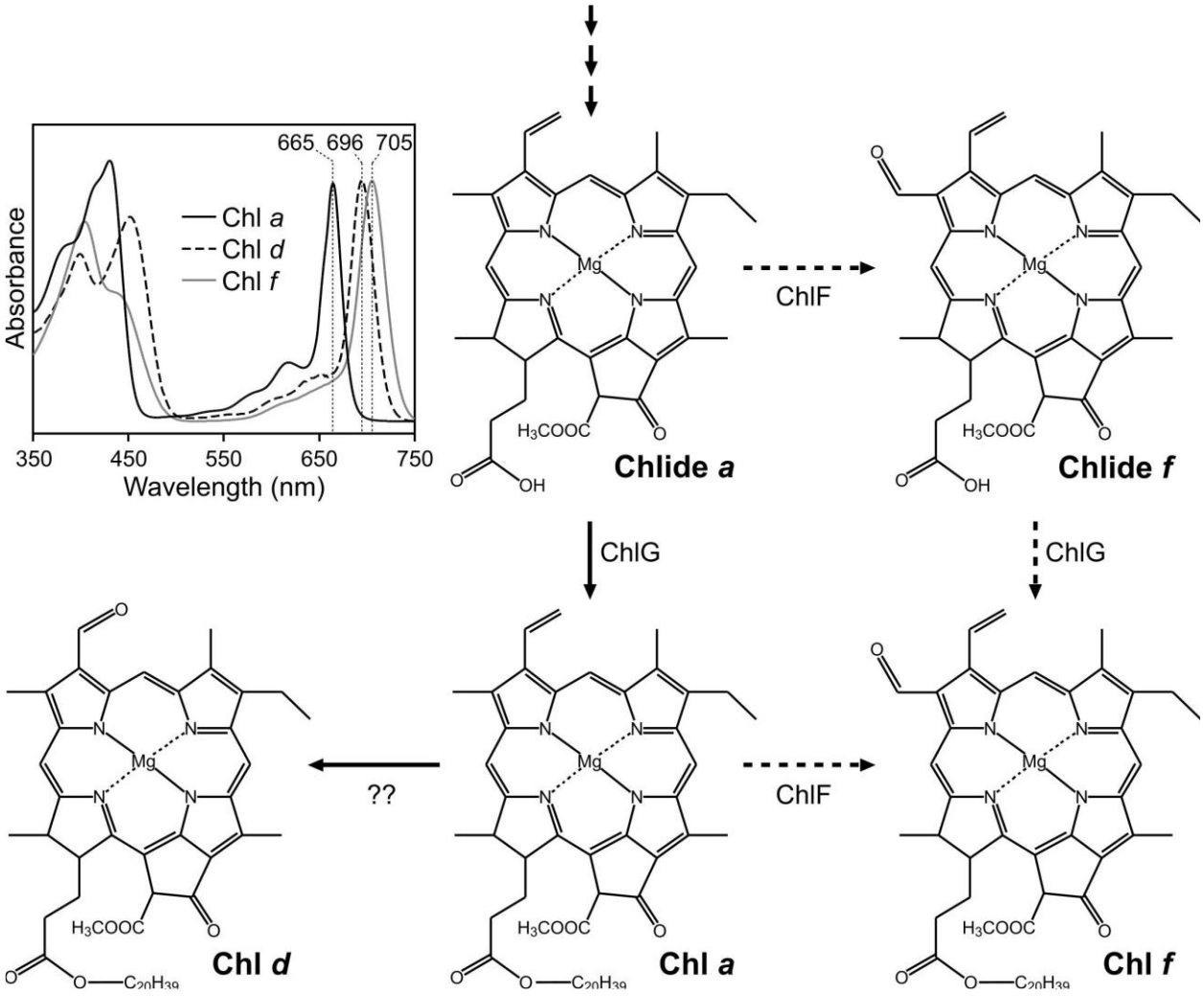


Fig. 1. Proposed biosynthetic pathways for Chl *a*, Chl *d*, and Chl *f*. The question marks indicate that the enzyme(s) that form Chl *d* from Chl *a* is/are not yet known, although this reaction pathway has been validated (*50*). ChlF converts either Chlide *a* or Chl *a* into Chlide *f* or Chl *f*. It is assumed that ChlG adds the esterifying alcohol phytol if Chlide *f* is formed as an intermediate. The panel at the upper left shows in-line absorption spectra for Chl *a*, Chl *d*, and Chl *f* with Q_y absorption maxima at 665, 696, and 705 nm, respectively.

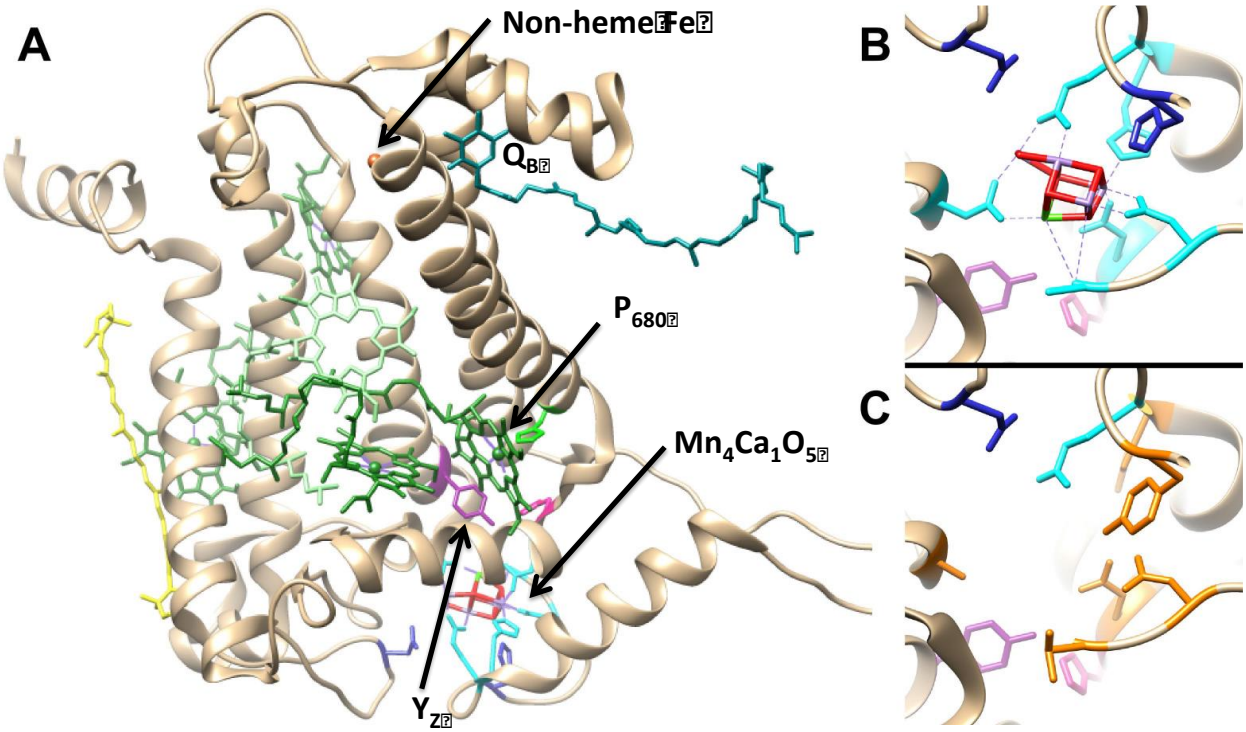


Fig. 2. Structural model of Photosystem II core components from *Thermosynechococcus vulcanus* (PDB 3WU2). A. PsbA (Chain A) is shown in ribbon format in tan**,** three Chl *a* molecules and one pheophytin *a* molecule are shown in green; plastoquinone Q_B is shown in teal; non-heme iron atom is shown in orange; tyrosine Y_Z is shown in magenta; β-carotene is shown in yellow; and the water-oxidizing $Mn_4Ca_1O_5$ cluster is shown in (lilac, green, and red). Amino acid side chain ligands to the $Mn_4Ca_1O_5$ cluster are shown in aqua, and second-tier amino acids in the vicinity of the cluster are shown in blue. **B.** Enlargement of the $Mn_4Ca_1O_5$ cluster that now shows a histidine residue (pink) that may be involved in proton-coupled electron transfer in PSII. Within the cluster, manganese atoms are shown in lilac, the calcium atom is shown in green, and bridging oxygen atoms are shown in red. Ligands to the $Mn_4Ca_1O_5$ cluster are shown in aqua. **C.** Model of the sr-PsbA4 $Mn_4Ca_1O_5$ cluster-binding site of *Chl. fritschii* PCC 9212. Only one ligand, a glutamate residue (aqua), and the C-terminal carboxyl group ligands of a non-conserved threonine residue are retained. Amino acids are those that replace the conserved ligands to the $Mn_4Ca_1O_5$ cluster are shown in orange. These include a second tyrosine residue found in seven of thirteen PsbA4 sequences. Note that tyrosine Y_Z and the associated histidine residue are conserved and are shown in magenta and pink, respectively. The color coding used in this figure matches the color coding in the multiple sequence alignment file shown in Fig. S2. Images were created with the UCSF Chimera program (*51*), and residues were substituted with conformations of highest probability using the Dunbrack backbone-dependent rotamer library (*52*).

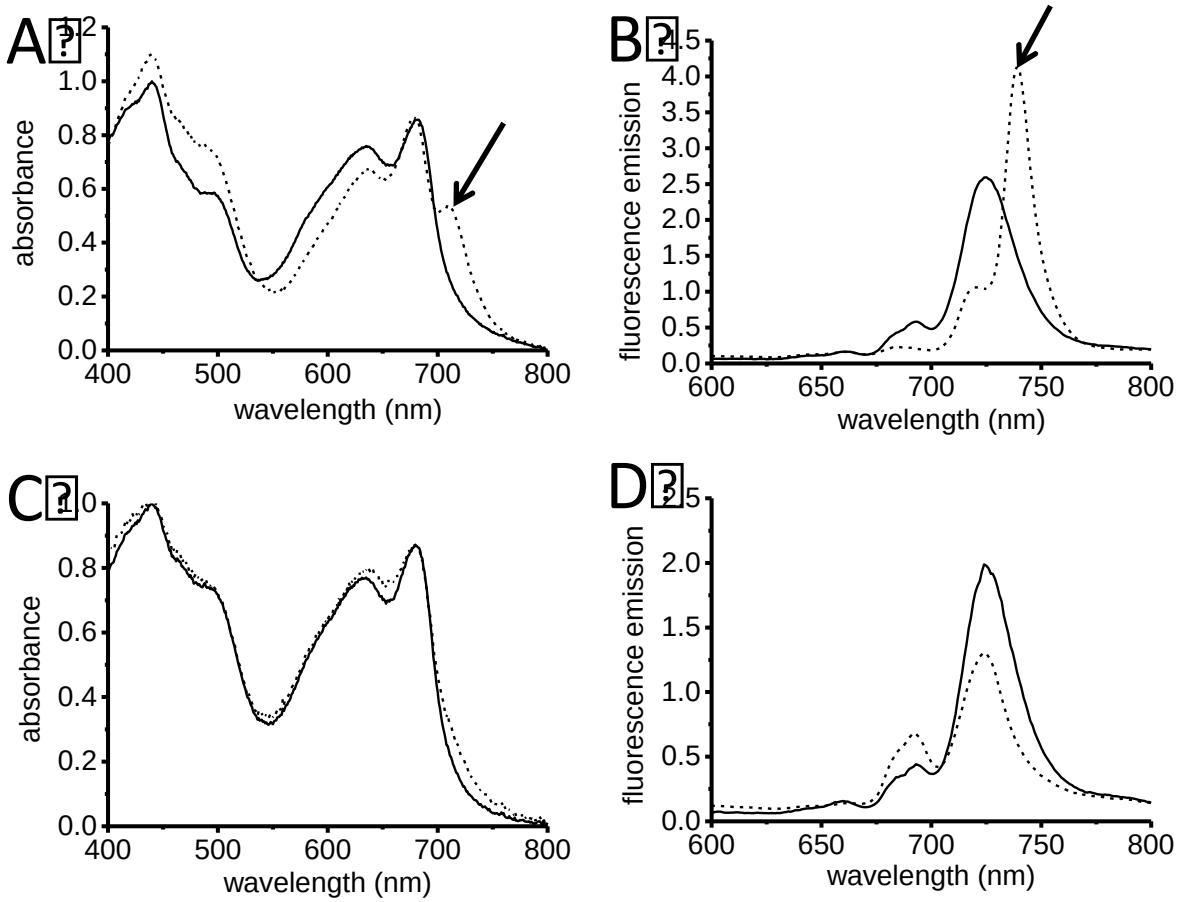


Fig. 3. The *psbA4* mutant of *Chl. fritschii* PCC 9212 lacks far-red absorbance and fluorescence emission. Whole-cell absorption spectra (**A, C**) and low-temperature fluorescence emission spectra (**B, D**) of wild-type (**A, B**) and *psbA4* mutant cells (**C, D**) of *Chl. fritschii* PCC 9212 grown in WL (solid lines) or FRL (dashed lines). The arrows in panels **A** and **B** point to long-wavelength absorption and fluorescence features due to Chl *f*.

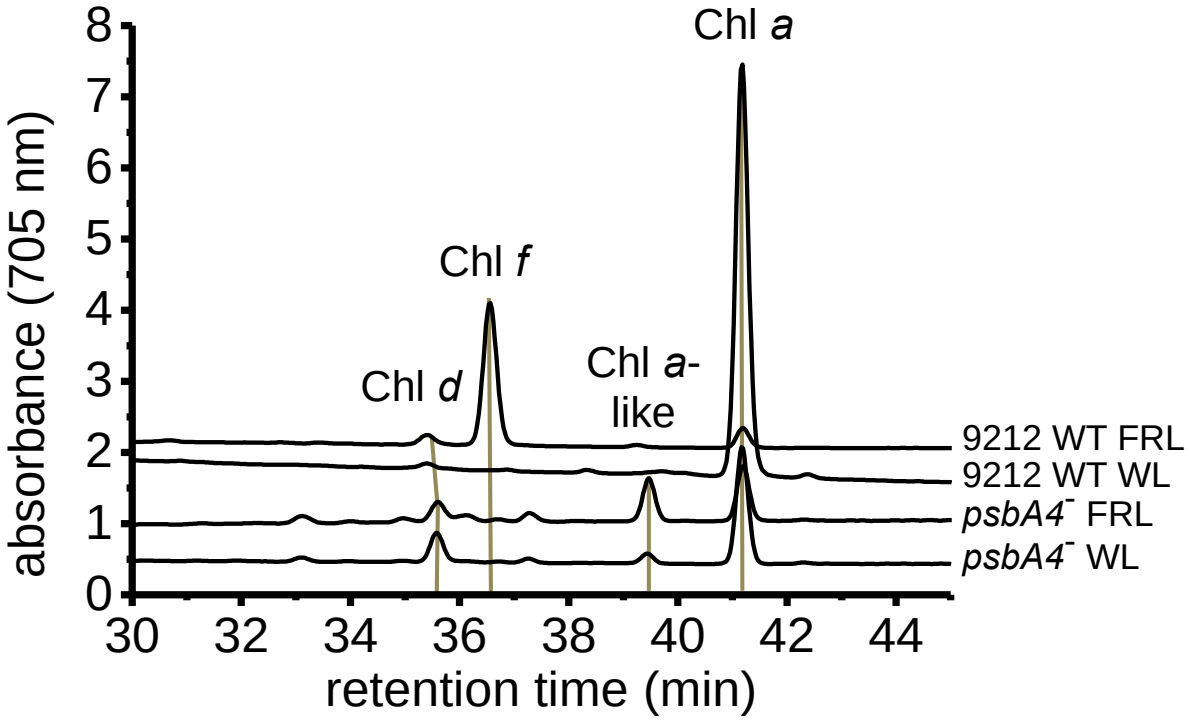


Fig. 4. The *psbA4* mutant of *Chl. fritschii* PCC 9212 is unable to synthesize Chl *f* but still produces Chl *d* and Chl *a*. Reversed-phase HPLC analysis using method MY1 (see Materials and Methods) of pigment extracts from wild-type and *psbA4* mutant cells of *Chl. fritschii* PCC 9212 grown in WL or FRL. Note that Chl *f* is present in the wild-type pigments extracted from cells grown in FRL but not in *psbA4* mutant cells. The peak labeled "Chl *a*-like" had the absorption spectrum of Chl *a* but had a different retention time.

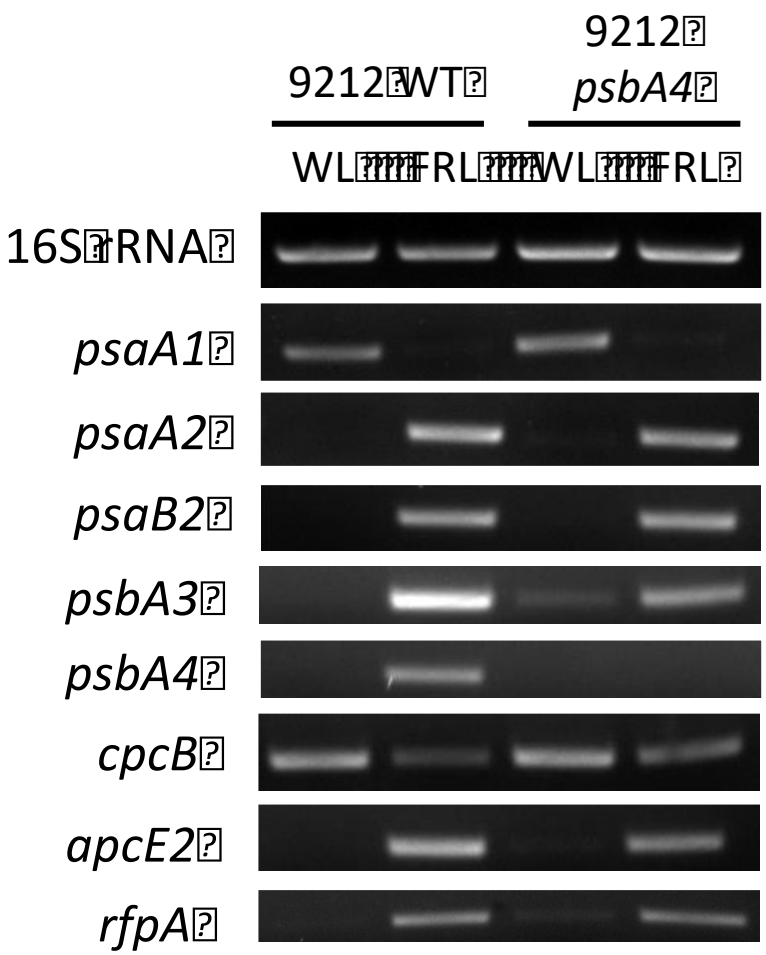


Fig. 5. Transcriptional regulation by RfpABC system is not impaired in the *psbA4* mutant of *Chl. fritschii* PCC 9212. RT-PCR analysis of transcripts for selected genes from the FaRLiP gene cluster (*psaA2, psaB2, psbA3, psbA4, apcE2,* and *rfpA*) from wild-type and *psbA4* mutant cells of *Chl. fritschii* PCC 9212 grown in WL or FRL. Paralogous genes from elsewhere in the genome are more highly expressed in WL (*psaA1, cpcB*). Note that no *psbA4* transcripts are observed in the *psbA4* mutant grown in FRL.

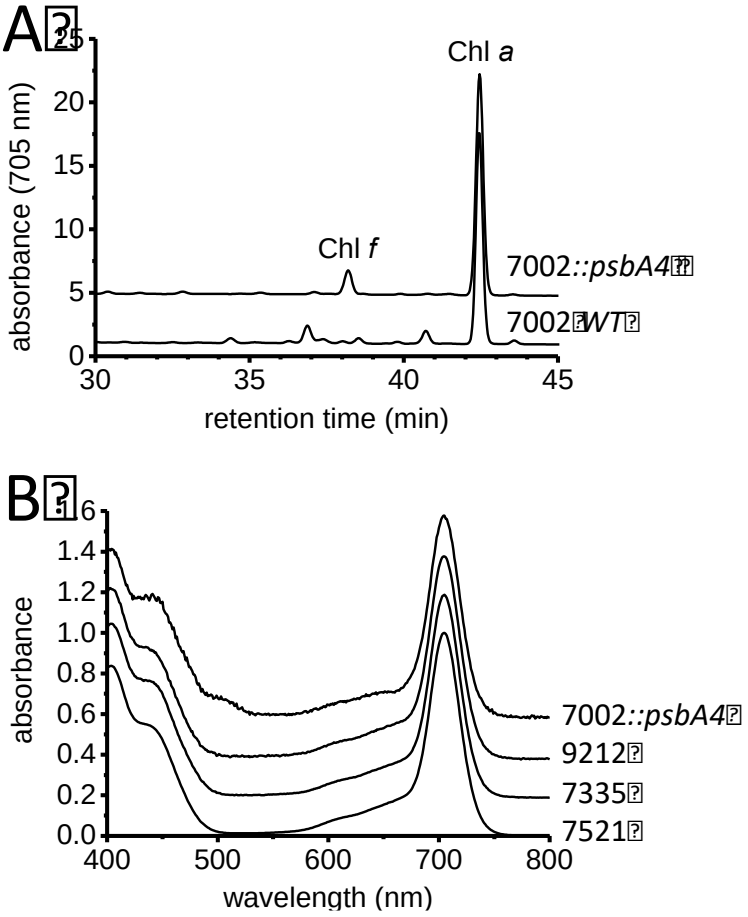


Fig. 6. Synthesis of Chl *f* in a *Synechococcus* 7002 strain expressing the *psbA4* gene of *Chl. fritschii* PCC 9212. A. Reversed-phase HPLC elution profile at 705 nm for pigments extracted from wild-type *Synechococcus* 7002 and a strain expressing the *psbA4* gene from *Chl. fritschii* PCC 9212 (7002*::psbA4*). The pigments were analyzed using method MY1 (see Materials and Methods). Chl *f* was 0.059% of the total Chl in the cells. **B.** In-line absorption spectra of Chl *f* from a *Synechococcus* 7002 strain expressing the *psbA4* gene from *Chl. fritschii* PCC 9212 (7002*::psbA4*). The absorption spectra of Chl *f* from *Chl. fritschii* PCC 9212 (9212), *Synechococcus* sp. PCC 7335 (7335), and *Fischerella thermalis* PCC 7521 (7521) were identical to Chl *f* from 7002*::psbA4*. Spectra were normalized to equal absorbance at 705 nm and were offset for clarity.

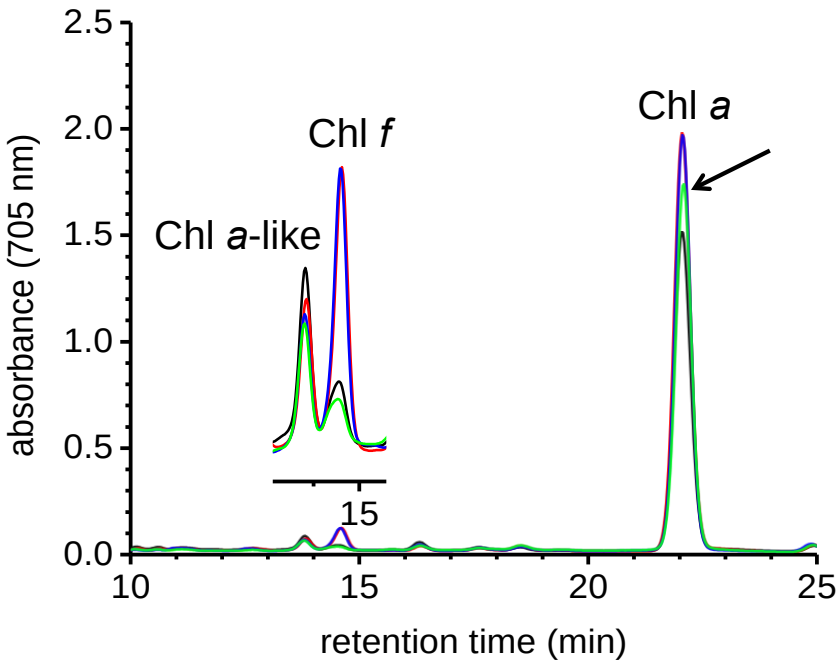


Fig. 7. Chl *f* synthesis, but not Chl *a* synthesis, requires light in *Synechococcus* 7002 expressing PsbA4/ChlF. HPLC elution profile of pigments extracted from *Synechococcus* 7002 expressing *chlF/psbA4* using method MY2 (see Materials and Methods for details). The elution positions for an unidentified Chl *a*-like pigment, Chl *f*, and Chl *a* are indicated. Cells were grown in white light for 24 h, which resulted in a basal level of Chl *f* synthesis (black line). These cells were used to inoculate cultures that were grown in medium A^+ containing 10 mM glycerol as a carbon/energy source. Cells were grown in darkness (green line), low-intensity WL (red line), or FRL (blue line). Cells grown in darkness for 3 days had a lower Chl *f* content than the cells had at time zero. However, cells incubated in the dark grew and synthesized new Chl *a* (arrow). Cells grown in FRL (blue line) and low WL (red line) also grew and synthesized more Chl *a* than cells grown in the dark. However, only cells grown in continuous FRL and low-intensity WL were able to synthesize Chl *f* (see inset). This experiment was performed twice and identical results were obtained.

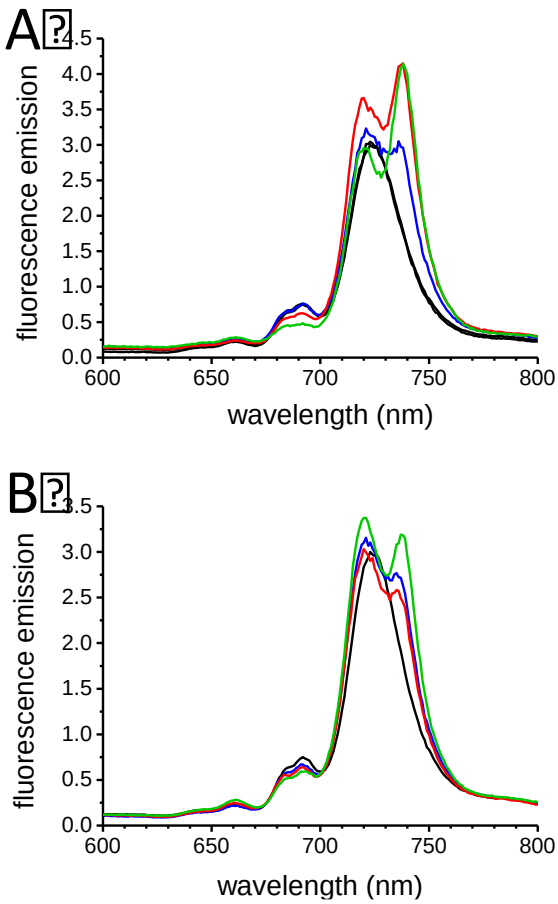


Fig. 8. Chl *f* synthesis in wild-type cells of *Chl. fritschii* PCC 9212 requires light. A. A starter culture was grown in WL in a medium containing 5 mM fructose, and at time zero, these cells were used to inoculate two cultures in the same medium. One was grown in continuous FRL and low-temperature fluorescence emission spectra were taken at 0 (black line), 24 (blue line), 36 (red line), and 48 h (green line). **B.** The second culture was incubated in FRL for 24 h, shifted to darkness for 12 hours, and then returned to FRL for 12 h. Low-temperature fluorescence emission spectra were taken at 0 (black line), after 24 h in FRL (blue line), after 12 h in the dark (36 h; red line), and after an additional 12-h incubation in FRL (48 h; green line). Chl *f* synthesis stopped during the 12-h dark period but resumed when cells were returned to FRL.

Supplementary Materials for

Light-dependent chlorophyll *f* synthase is a highly divergent paralog of PsbA of Photosystem II

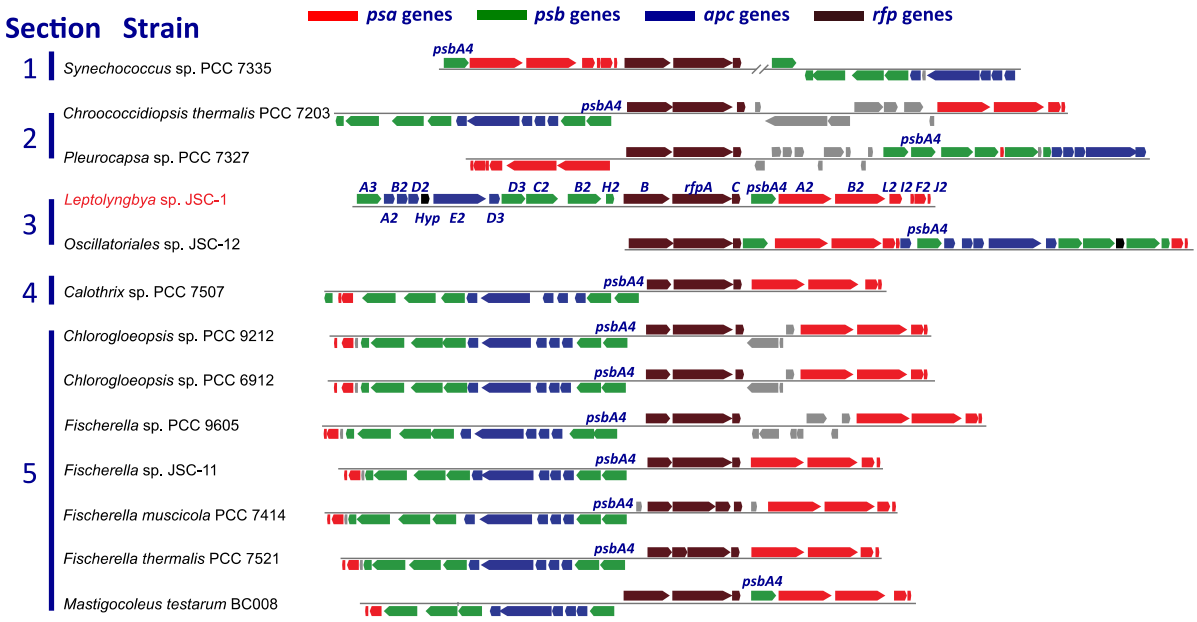
Ming-Yang Ho, Gaozhong Shen, Daniel P. Canniffe, Chi Zhao, and Donald A. Bryant*

*correspondence to: dab14@psu.edu

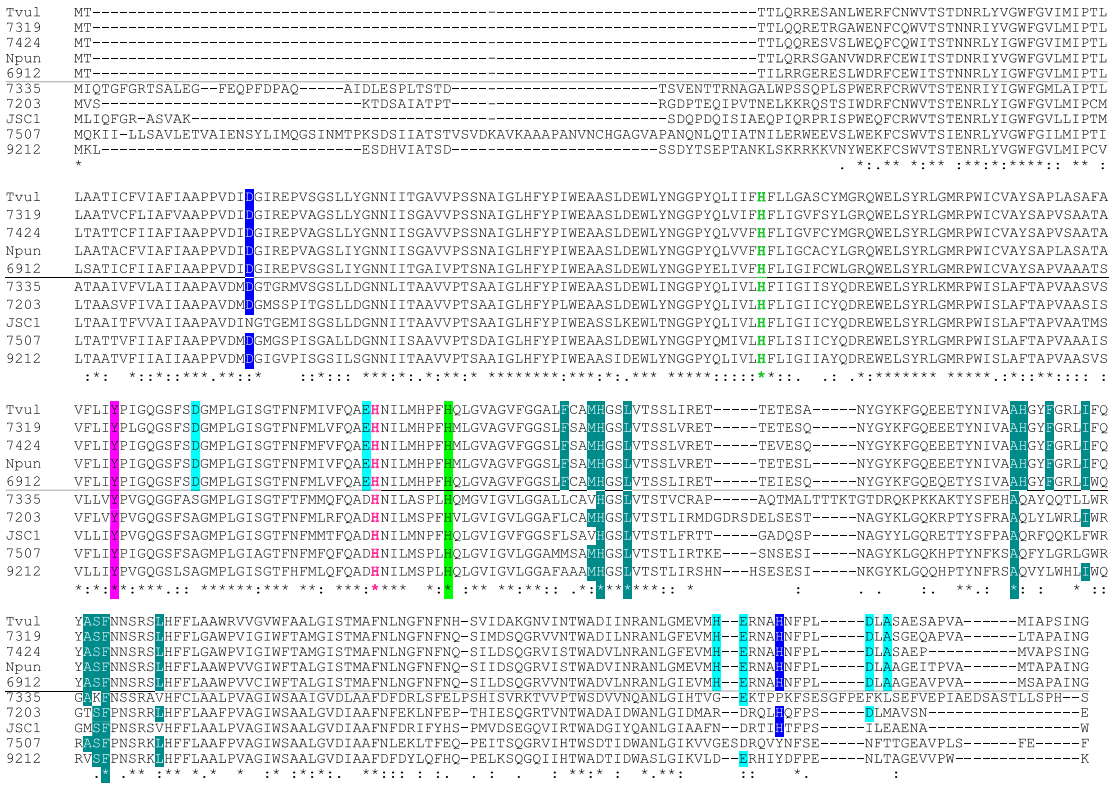
This PDF file includes:

Figures S1 – S8
Tables S1 – S2

Supplemental Figure S1

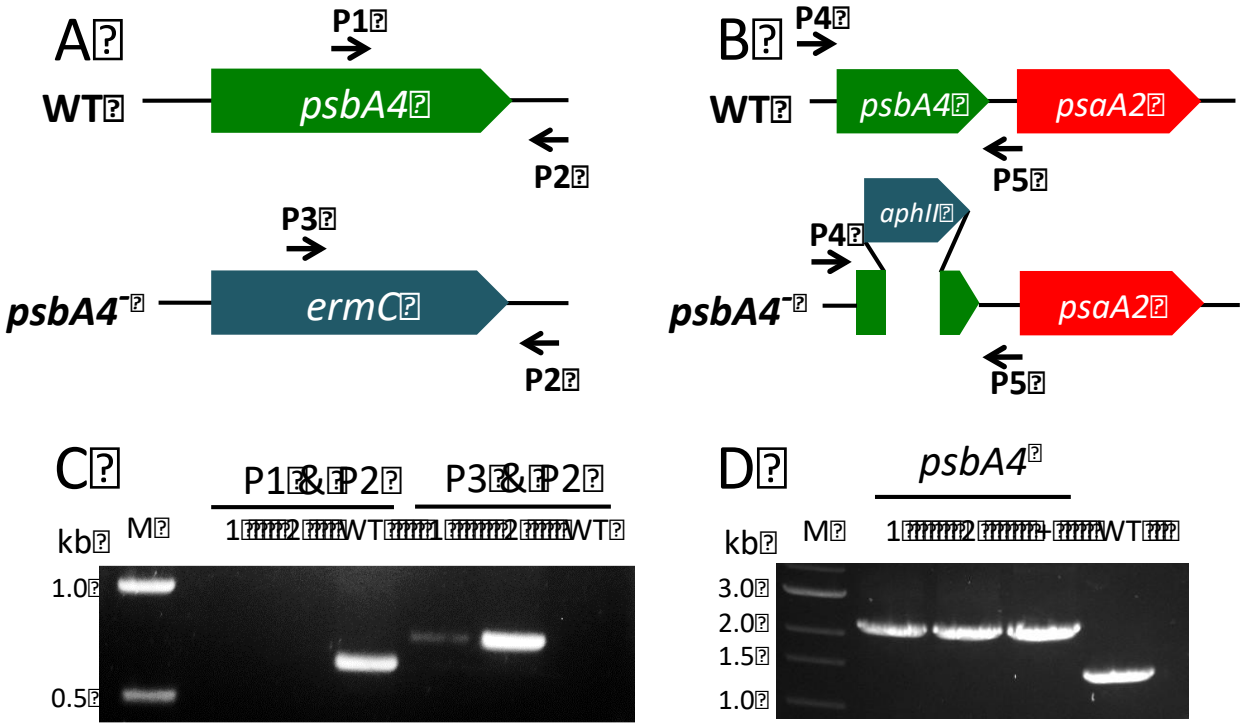


Supplemental Fig. S1. Comparison of FaRLiP gene clusters in the cyanobacterial strains indicated. Genes encoding core components of the phycobilisome are shown in blue, of PSI in red, of PSII in green. Regulatory genes *rfpA*, *rfpB*, and *rfpC* are shown in dark red-brown. The taxonomic sections for the various cyanobacteria that can perform FaRLiP are indicated at the left. The *psbA4* (*chlF*) gene in each cluster is indicated.

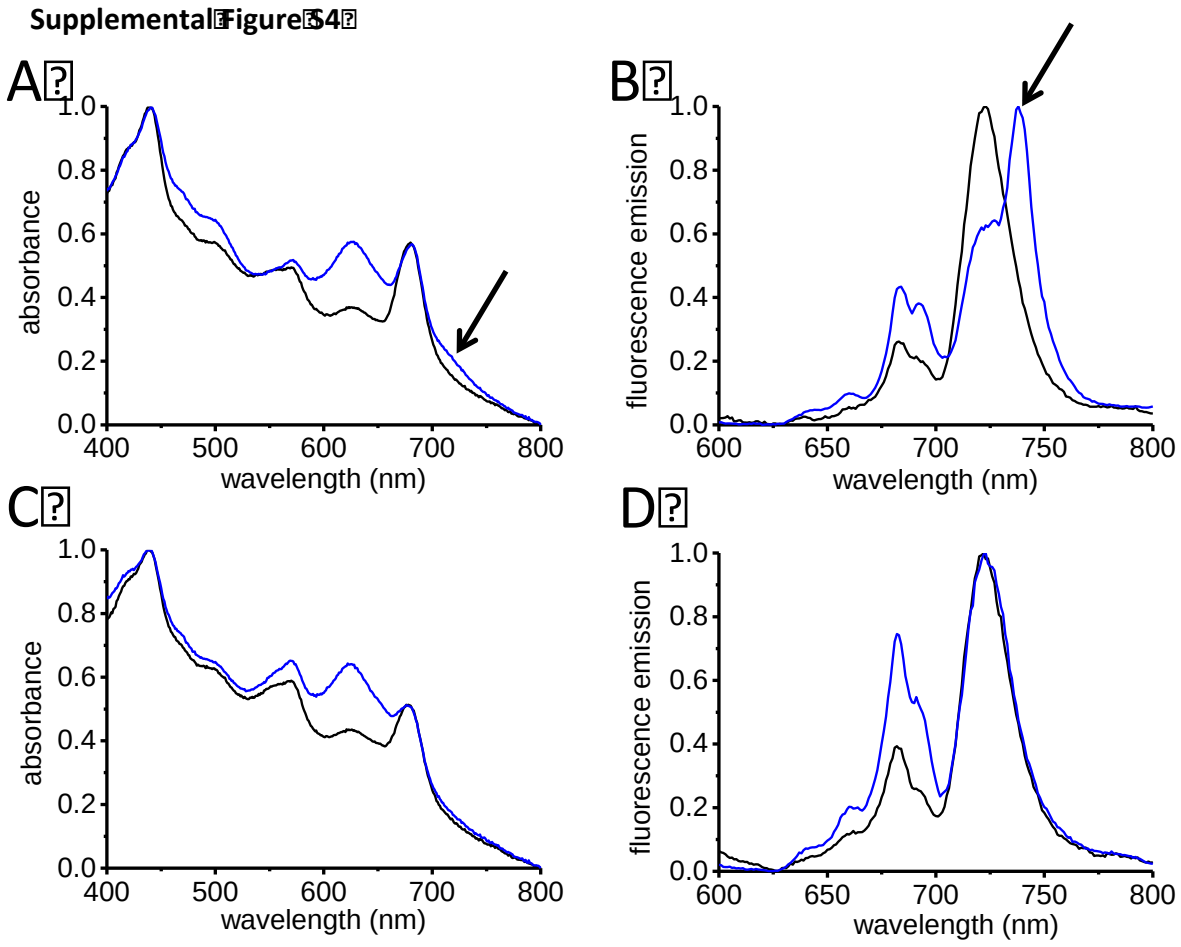


Supplemental Fig. S2. Multiple sequence alignment of selected conventional PsbA/D1 proteins with PsbA4/ChlF proteins, listed above and below the dividing line, respectively. Each of the five conventional PsbA and ChlF sequences are examples from taxonomic subsections I-V, listed in ascending order. Tyrosine Y_Z is shaded magenta; the histidine in close proximity to Y_Z, also believed to be involved in proton-coupled electron transfer, is shown in pink font; ligands to the $Mn_4Ca_1O_5$ oxygen-evolving center are shaded aqua; the histidine residue ligating a P680 Chl *a* is shaded green; the histidine residue ligating an additional Chl *a* is shown in green outline font; residues in the second coordination sphere of the $Mn_4Ca_1O_5$ cluster are shaded blue; and residues composing the Q_B binding site are shaded teal. Note that most of the residues ligating the $Mn_4Ca_1O_5$ cluster are not conserved in PsbA4/ChlF. Conserved, highly similar, and similar residues are denoted by asterisks, colons, and periods, respectively. Tvul, *Thermosynechococcus vulcanus*; 7319, *Pleurocapsa* sp. PCC 7319; 7424, *Cyanothece* sp. PCC 7424; Npun, *Nostoc punctiforme*; 6912, *Chl. fritschii* PCC 6912; 7335, *Synechococcus* sp. PCC 7335; 7203, *Chroococcidiopsis thermalis* PCC 7203; JSC1, *Leptolynbya* sp. JSC-1; 7507, *Calothrix* sp. PCC 7507; 9212, *Chl. fritschii* PCC 9212.

Supplemental Figure S3

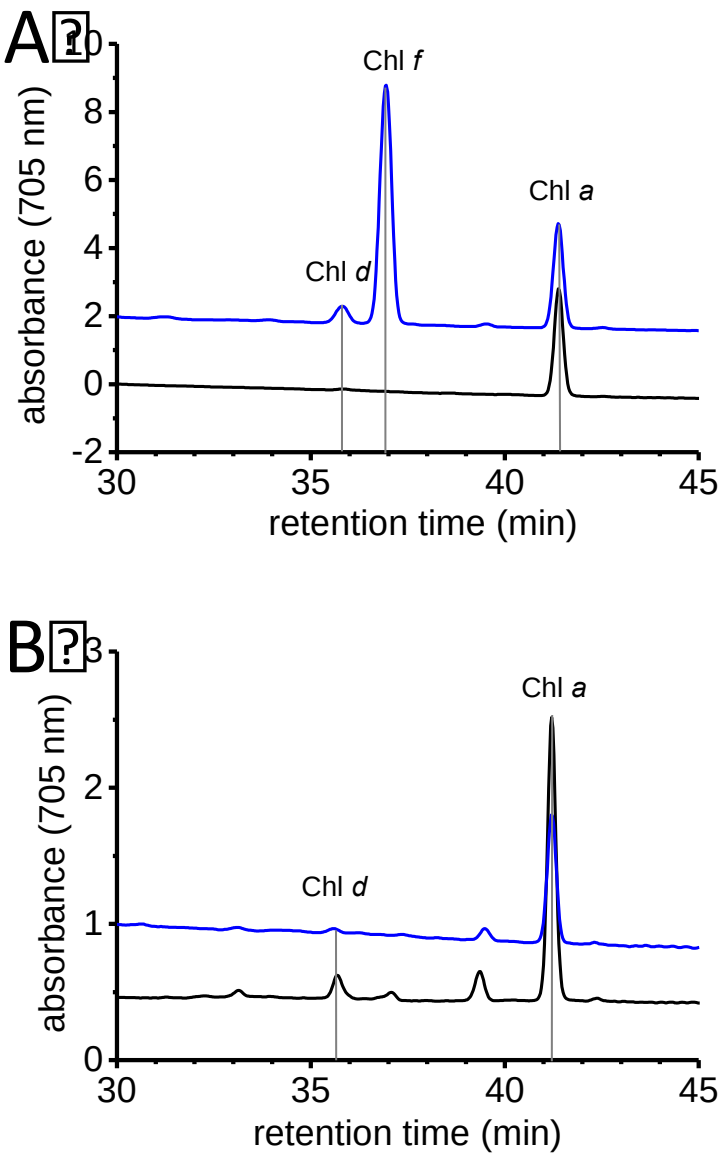


Supplemental Fig. S3. Construction and validation of *psbA4* mutants in *Chl. fritschii* PCC 9212 and *Synechococcus* sp. PCC 7335. Schemes showing the construction of mutants deleting *psbA4* genes in *Chl. fritschii* PCC 9212 (A) or *Synechococcus* sp. PCC 7335 (B). The small numbered arrows identify oligonucleotide primers used to verify the constructions by PCR (see Table S2). Agarose gel electrophoretic analysis of PCR amplicons produced using the primer combinations indicated and template DNAs from two transconjugants or the wild-type (WT) strain of in *Chl. fritschii* PCC 9212 (**C**) or *Synechococcus* sp. PCC 7335 (**D**). PCR amplicons were also isolated and verified by DNA sequencing. M, size markers. WT, wild-type DNA template; 1 and 2, independent transconjugants. +, positive control (cargo plasmid used for conjugation). The primer pairs used for verification of the mutations in *Chl. fritschii* PCC 9212 are indicated above in Panel **C**.



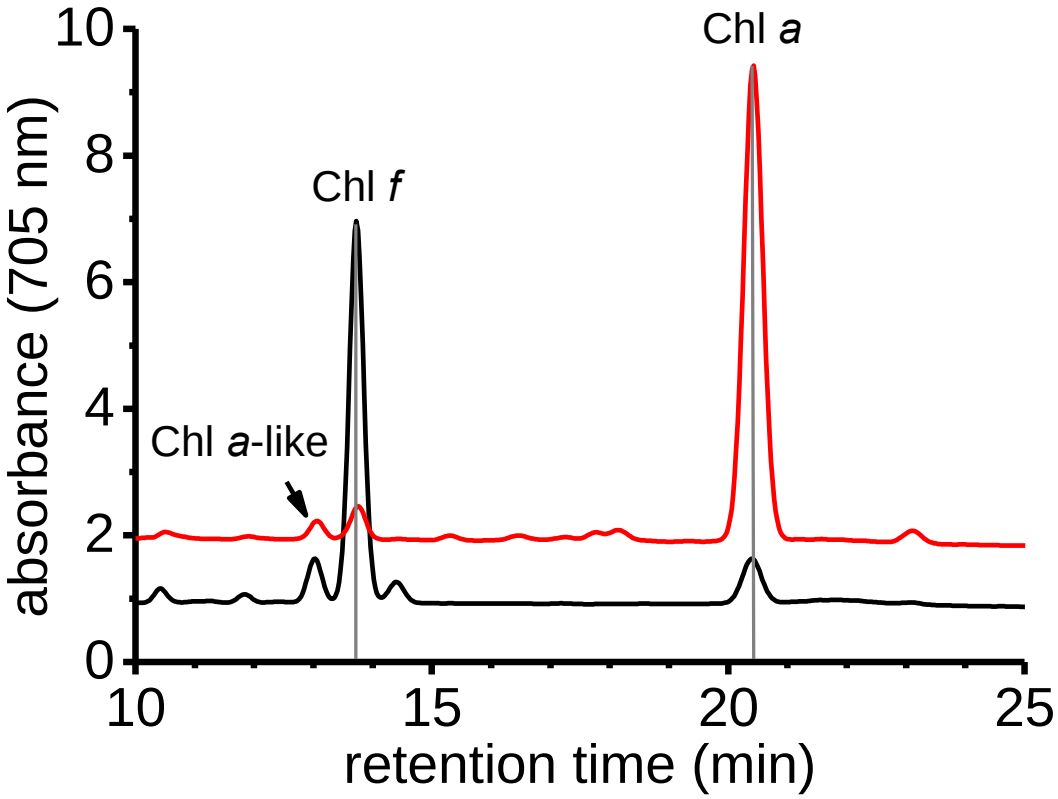
Supplemental Fig. S4. A *psbA4* mutant of of *Synechococcus* sp. PCC 7335 lacks absorption and fluorescence features characteristic of Chl *f*. Whole-cell absorption spectra (**A, C**) and low-temperature fluorescence emission spectra (**B, D**) of wild-type (**A, B**) and *psbA4* mutant cells (**C, D**) of *Synechococcus* sp. PCC 7335 grown in WL (black lines) or FRL (blue lines). The arrows in panels **A** and **B** point to long-wavelength absorption and fluorescence features due to Chl *f*.

Supplemental Figure S5



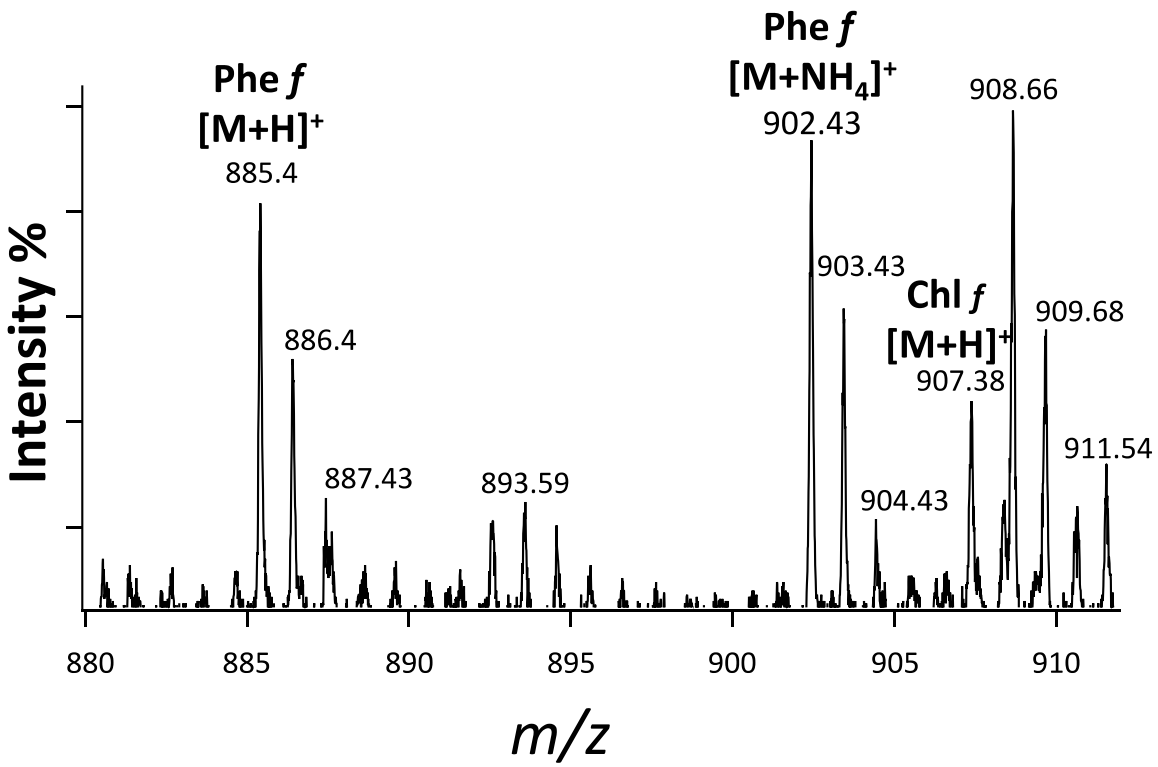
Supplemental Fig. S5. A *psbA4* mutant of *Synechococcus* sp. PCC 7335 is no longer able to synthesize Chl *f* when cells are grown in FRL. Reversed-phase HPLC analysis of pigment extracts from wild-type (**A**) and *psbA4* mutant cells (**B**) of *Synechococcus* sp. PCC 7335 grown in WL (black lines) or FRL (blue lines). Pigments were analyzed using method MY1 (see **Materials and Methods**). Note that Chl *f* is present in the wild-type extract from cells grown in FRL but not in the *psbA4* mutant cells grown in FRL.

Supplemental Figure S6

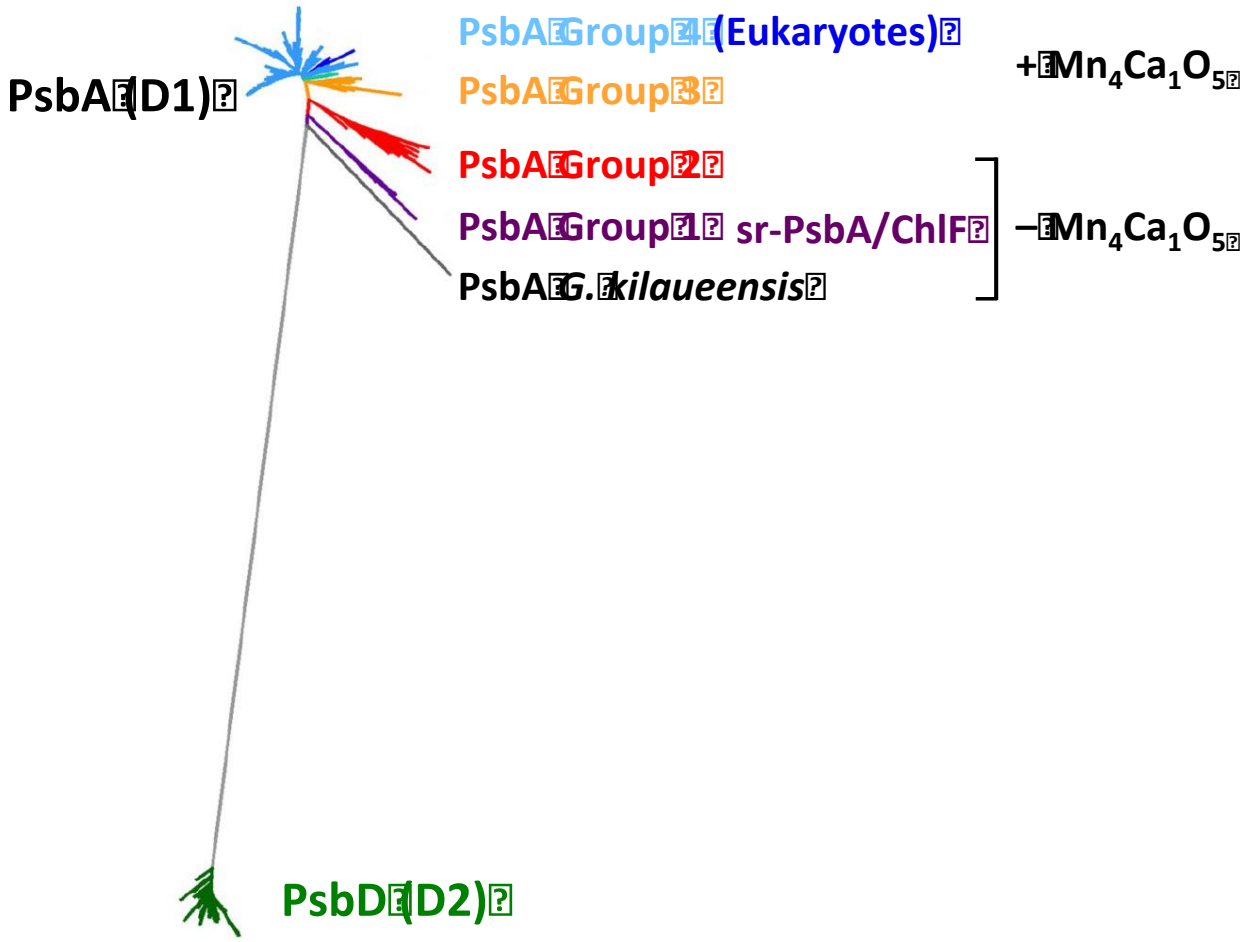


Supplemental Fig. S6. Chl *f* synthesized by *Synechococcus* 7002 cells expressing *psbA4/chlF* has the elution time as authentic Chl *f* produced in *Chl. fritschii* PCC 9212. Comparison of reversed-phase HPLC elution profiles monitored at 705 nm of pigment extracts from wild-type cells of *Chl. fritschii* PCC 9212 grown in FRL (black line) and *Synechococcus* 7002 cells expressing the *psbA4* gene from *Chl. fritschii* PCC 9212 (red line). Note that the retention time of Chl *f* from these two sources is identical (chromatography using method MY2 (see **Materials and Methods**) was performed on the same day). The peak labeled "Chl *a*-like" is an unidentified pigment with an absorption spectrum identical to Chl *a*.

Supplemental Figure S7



Supplemental Fig. S7. **Mass spectrophotometric analysis of purified Chl *f* from *Synechococcus* 7002 expressing PsbA4.** The Chl *f* was purified using semi-preparative HPLC chromatography by method MY2, and the compound was then injected with a mobile phase containing formic acid to convert Chl *f* partially to Phe *f*. For details, see **Materials and Methods**.



Supplemental Fig. S8. Maximum likelihood phylogenetic tree of PsbA sequences with PsbD (D2) as outgroup. This figure is adapted from Fig. 1 of Cardona *et al.* (*28*). The PsbA sequences from *Gloeobacter kilaueensis*, Group 1, and Group 2 lack most of the ligands required for binding the water-oxidizing $Mn_4Ca_1O_5$ cluster and exhibit changes in the plastoquinone binding site. Murray (*27*) described Group 1 sequences (purple) as "super-rogue" PsbA sequences and Group 2 sequences (red) as "rogue" PsbA sequences. All Group 1 sequences occur in FaRLiP gene clusters and were shown here to encode Chl *f* synthase (ChlF). The PsbA sequences of Group 3 (orange) and Group 4 (blue) have all of the sequence and structural features required for water oxidation. Some group 3 sequences are expressed under microoxic conditions. Eukaryotic PsbA sequences (dark blue) form a distinct subclade within the Group 4 sequences. Cardona et al. (*28*) have suggested that the sequences of *G. kilaueensis*, Group 1, and Group 2 evolved prior to the PsbA sequences of Groups 3 and 4 that can support water oxidation. See main text for additional details.

Supplemental Table S1

In-solution tryptic digestion and LC-MS-MS identification of proteins in PSI complexes. See **Materials and Methods** for details.

Trimeric PSI complexes from *Chl. fritschii* PCC 9212 WT (WL)

Accession	Score (-10lgP)	Unique Peptides	Protein Name
2550829734	362.11	53	PsaA1
2550829735	338.59	24	PsaB1
2550834188	302.06	39	PsaL1
2550834190	290.54	41	PsaF1
2550833654	285.49	23	PsaE
2550831352	261.99	26	PsaD
2550832806	226.86	13	hypothetical
2550833343	165.01	7	PsaC
2550831271	142.33	3	hypothetical
2550828414	129.32	5	hypothetical
2550832263	112.61	4	PsaX
2550830136	105.87	3	hypothetical
2550829150	63.15	1	PsaK
2550828976	54.7	1	hypothetical
2550834189	54.08	3	PsaJ1
2550832220	53.5	1	hypothetical
2550833592	43.45	1	hypothetical
2550829987	29.3	1	PsaI1
2550831528	22.05	1	hypothetical

Trimeric PSI complexes from *Chl. fritschii* PCC 9212 WT (FRL)

Accession	Score (-10lgP)	Unique Peptides	Protein Name
2550828630	321.18	24	PsaA2
2550828629	272.13	13	PsaB2
2550829734	264.87	11	PsaA1
2550829735	241.17	10	PsaB1
2550833654	230.29	8	PsaE
2550834190	229.74	10	PsaF1
2550828628	219.04	7	PsaL2
2550831352	211.12	11	PsaD
2550832645	211.12	7	hypothetical
2550828650	202.62	12	PsaF2

2550831271	201.76	6	hypothetical
2550834188	201.5	5	PsaL1
2550830636	199.36	3	PsaB3
2550828251	184.83	8	hypothetical
2550833343	168.04	5	PsaC
2550830218	163.71	4	hypothetical
2550832159	159.95	4	Ycf66
2550831214	135.87	3	Ycf4
2550833810	135.75	3	hypothetical
2550831709	121.39	4	RubA
2550829447	111.5	3	thioredoxin
2550833287	110.6	2	hypothetical
2550832263	109.39	4	PsaX
2550827920	99.76	3	hypothetical
2550834189	35.59	1	PsaJ1
2550830755	34.89	1	hypothetical
2550829987	28.85	1	PsaM
2550830775	26.83	1	hypothetical
2550827818	25.8	1	hypothetical

Trimeric PSI complexes from *Chl. fritschii* PCC 9212 *psbA4* mutant (WL)

Accession	Score (-10lgP)	#Unique	Protein Name
2550829735	332.93	25	PsaB1
2550829734	317.62	41	PsaA1
2550834188	285.9	37	PsaL1
2550833654	269.15	24	PsaE
2550831352	251.16	28	PsaD
2550834190	215.55	23	PsaF1
2550832806	181.36	8	hypothetical
2550832759	111.76	5	hypothetical
2550829150	95.93	4	PsaK
2550832263	76.31	2	PsaX
2550833287	64.44	2	hypothetical
2550833343	60.86	2	PsaC
2550833810	53.58	2	hypothetical
2550828414	36.78	1	hypothetical
2550834189	35.32	1	PsaJ1
2550830093	34.92	1	hypothetical
2550829987	31.11	1	PsaM
2550833890	25.44	1	hypothetical

Trimeric PSI complexes from *Chl. fritschii* PCC 9212 *psbA4* mutant (FRL)

Accession	Score (-10lgP)	#Unique peptides	Protein Name
2550829735	315.07	16	PsaB1
2550829734	301.6	22	PsaA1
2550834188	276.25	26	PsaL1
2550834190	274.59	26	PsaF1
2550833654	247.12	14	PsaE
2550831352	237.7	19	PsaD
2550830636	225.03	2	PsaB3
2550833631	210.22	9	hypothetical
2550828630	199.74	2	PsaA2
2550832805	185.88	9	hypothetical
2550833343	178.3	6	PsaC
2550831803	175.01	7	hypothetical
2550828812	130.51	3	hypothetical
2550828629	122.32	1	PsaB2
2550830391	119.43	3	hypothetical
2550831709	87.02	2	RubA
2550831683	84.55	2	hypothetical
2550829150	55.96	1	PsaK
2550831572	53.69	1	hypothetical
2550831214	48.72	1	Ycf4
2550832263	42.62	0	PsaX
2550834189	40.76	2	PsaJ1
2550833905	38.82	1	hypothetical

Supplemental Table S2

Oligonucleotide primers used for PCR, RT-PCR, and strain construction.

Km-1 SalI	5'-TTATGTCGACCACGCTGCCGCAAGCACTCA-3'
Km-2 BamHI	5'-ATAGGATCCGGTGGGCGAAGAACTCC-3'
7335 *psbA4* up-1 (SacI)	5'-TGACGAGCTCTAAACTCATCGCGAATGCTG-3'
7335 *psbA4* up-2 (SalI)	5'-GTCAGTCGACCCGTGGCTAAGGTTGGAATA-3'
7335 *psbA4* down-1 (BamHI)	5'-ATATGGATCCGATGGCCCTAACGACGACTA-3'
7335 *psbA4* down-2 (XhoI)	5'-TAGTCTCGAGAGCCGTGGCAAGAGAGTAAA-3'
Em-1 SacI	5'-GCGCGAGCTCGCGTGCTATAATTATACTAATTTTATAA GGAGGAAAAAAT-3'
Em-2 BamHI	5'-TTAAGGATCCTGAGTGAGCTGATACCGCTCGCCGCAG-3'
9212 *psbA4* up-1 (XhoI)	5'-TTAACTCGAGCTCCCGCTGTTGTCTTTAACAGCAGAAAT-3'
9212 *psbA4* up-2 (SacI)	5'-GGTTGAGCTCGCTTCATACTTAATACATCCTTACGGAAT-3'
9212 *psbA4* down-1 (BglII)	5'-GGTTAGATCTCAGCAGGTGAAGTAGTGCCTTGGAAGT-3'
9212 *psbA4* down-2 (BamHI)	5'-TTAAGGATCCGCTACTTCGGCAGCGTCTTCT-3'
P1: 9212 *psbA4*-i_F	5'-TATCTGCGGGAATGCCTTTA-3'
P2: 9212 *psbA4*-d_R	5'-GTTGCAAAACCTGTCCCAAG-3'
P3: Em-i_F	5'-TTGTTGATCACGATAATTTCCAA-3'
P4: 7335 *psbA4*-u_F	5'-GGAAGCAGCGTCTCCATTAG-3'
P5: 7335 *psbA4*-d_R	5'-CGAAGACCAATTACCAAAGACTG-3'
9212 16S RT-1	5'-GATACTAGGCGTTGCGTGTATCGACCCAC-3'
9212 16S RT-2	5'-GTTGCAGCCTGCGATCTGAACTGAG-3'
9212 *psaA1* RT-1	5'-CAAAGACTTGTTGACCGAGCTGTATCCTAGC-3'
9212 *psaA1* RT-2	5'-TGAAGATGGCAGCGTGAGCAGCACC-3'
9212 *psbB1* RT-1	5'-GTTGGCGATCCCAATTACAATCCAGGTTTC-3'
9212 *psbB1* RT-2	5'-AAGAATACAGCAGCAATACTGCTAGAAAGCACGG-3'
9212 *cpcB* RT-1	5'-ATGTTAGACGCATTTGCCAAGGTGGTTTC-3'
9212 *cpcB* RT-2	5'-GCGCGATCGAAGTAGCTAGCCAATTCAGAC-3'
9212 *psaA2* RT-1	5'-TCGTACCAATTGGGGTATCGGTCACAGCAT-3'
9212 *psaA2* RT-2	5'-TATCTTGAGGACGACCAAAAGCCCGCAT-3'
9212 *psaB2* RT-1	5'-CAAGGTGCAGGTACAGCAATTCTCACATTC-3'
9212 *psaB2* RT-2	5'-CGCTCTAAGACATTACCTTTATTCTGTTCTGGGTC-3'
9212 *psbA4* RT-1	5'-CAAGTGCCATAAATAAACTTGGGCAGACCT-3'
9212 *psbA4* RT-2	5'-ATCCTATTTGGGAAGCAGCTTCCATTGATGAG-3'
9212 *psbA3* RT-1	5'-GTAACCATGGGCTGCAACAATGTTATAGGT-3'
9212 *psbA3* RT-2	5'-CGATTGGATTGCACTTTTACCCAATTTG-3'
9212 *apcE2* RT-1	5'-GTCCACCCTTAACTTGGGATTCCAATTCAT-3'
9212 *apcE2* RT-2	5'-GCCGACGAGAATGAAACGGTCTATGCG-3'
9212 *rfpA* RT-1	5'-AGGAATACCTCGCCCTTGAT-3'
9212 *rfpA* RT-2	5'-TGACTTCAATTCGCAGTGCT-3'
9212 *psbA4* pAQ1-1 (NcoINdeI)	5'-GCATCCATGGCATATGAAGCTAGAGTCAGATCA-3'
9212 *psbA4* pAQ1-2 (BamHI)	5'-AGTCGGATCCCTACTTCCAAGGCACTACTT-3'