Chlorophyll *f* synthesis by a super-rogue photosystem II complex

Joko P. Trinugroho[1†], Martina Bečková[2†], Shengxi Shao[1†], Jianfeng Yu[1], Ziyu Zhao[1], James W. Murray[1], Roman Sobotka[2], Josef Komenda[2] and Peter J. Nixon[1*]

[1]Sir Ernst Chain Building-Wolfson Laboratories, Department of Life Sciences, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK.

[2]Laboratory of Photosynthesis, Centre Algatech, Institute of Microbiology, Czech Academy of Sciences, Třeboň, Czech Republic.

*Correspondence to: p.nixon@imperial.ac.uk

† equal contributions

Certain cyanobacteria synthesize chlorophyll (Chl) molecules (Chl *d* and Chl *f*) that absorb in the far-red region of the solar spectrum thereby extending the spectral range of photosynthetically active radiation[1,2] . The synthesis and introduction of these far-red Chls into the photosynthetic apparatus of plants might improve the efficiency of oxygenic photosynthesis, especially in far-red enriched environments, such as in the lower regions of the canopy[3]. Production of Chl *f* requires the ChlF subunit, also known as PsbA4[4] or super-rogue D1[5], a paralog of the D1 subunit of photosystem II (PSII) which together with D2 binds co-factors involved in the light-driven oxidation of water. Current ideas suggest that ChlF oxidizes Chl *a* to Chl *f* in a homodimeric ChlF reaction center (RC) complex and represents a missing link in the evolution of the heterodimeric D1/D2 RC of PSII[4,6]. However, unambiguous biochemical support for this proposal is lacking. Here we show that ChlF can substitute for D1 to form modified PSII complexes capable of producing Chl *f*. Remarkably mutation of just two residues in D1 converts oxygen-evolving PSII into a Chl *f* synthase. Overall, we have identified a new class of PSII

complex, which we term 'super-rogue' PSII, with an unexpected role in pigment biosynthesis rather than water oxidation.

There is a pressing need to improve crop productivity to meet the global demand for food which is predicted to double between 2005 and 2050[7]. One promising approach is to enhance the efficiency of photosynthesis, which uses the energy of sunlight to drive the synthesis of biomass from carbon dioxide[8,9]. Plants rely on two types of Chl pigment (Chl *a* and Chl *b*), plus accessory pigments, to absorb sunlight, which restricts photosynthesis to the visible region of the spectrum (wavelengths between 400-700 nm). Capturing light in the far-red region of the spectrum (700 - 750 nm) using far-red pigments has the potential to increase the number of photons available for photosynthesis by up to 19%[3].

Chl *f* synthesis involves the oxidation of a methyl group to a formyl group (Supplementary Fig. 1) but the nature of the cyanobacterial Chl *f* synthase enzyme is poorly understood. Two lines of evidence have indicated a role for ChlF in Chl *f* synthesis: inactivation of *chlF* blocked Chl *f* synthesis in native Chl *f*-producing cyanobacteria and heterologous expression of just ChlF led to light-dependent production of Chl *f* in the cyanobacterium *Synechococcus* sp. 7002, which is normally unable to make Chl *f* [4]. Although ChlF lacks the key amino-acid residues in D1 that bind the Mn_4CaO_5 metal cluster involved in water oxidation, making it unlikely to play a role in oxygen evolution, it does potentially retain a redox-active tyrosine on the donor side of the complex and residues that bind the Chl *a* and pheophytin *a* molecules involved in light-induced charge separation[5]. Whether ChlF forms a minimalistic homodimeric RC complex[4] or still binds to D2 and assembles into a larger multi-subunit PSII core complex is unresolved.

Analysis of FLAG-tagged ChlF complexes

To purify and characterize the Chl *f* synthase, we first expressed an N-terminal 3xFLAG-tagged derivative of ChlF (FLAG-ChlF) encoded by *Chroococcidiopsis thermalis* PCC 7203 in the

cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter Syn6803) lacking the D1 subunit and unable to synthesize Chl *f* (Supplementary Fig. 2). Analysis of whole-cell pigments by High Performance Liquid Chromatography (HPLC) confirmed that expression of ChlF led to the production of Chl *f* at levels of 0.24% of Chl *a* (Supplementary Table 1), in line with previous reports[4]. The identity of Chl *f* is apparent from its typical HPLC elution time and red-shifted absorption spectrum, which were both identical to those of Chl *f* separated from far red-light-treated cells of *C. thermalis* (Supplementary Fig. 3a,b). FLAG-ChlF was immunopurified from detergent-solubilised membranes using an anti-FLAG resin and analyzed by 2D gel electrophoresis involving clear native (CN) gel electrophoresis in the first dimension followed by denaturing SDS-PAGE in the second (Fig. 1a). Immunoblotting experiments (Supplementary Fig. 4) revealed that FLAG-ChlF was found mainly in a monomeric PSII-like core complex, designated PSII(1)', containing the CP47, CP43 and D2 Chl-binding core subunits plus low molecular mass subunits of PSII and in a smaller PSII-like sub-complex, termed RC47', that had lost CP43 during electrophoresis. The Psb27 subunit, which is diagnostic of non-oxygen-evolving PSII complexes[10], was detected, but not the extrinsic lumenal subunits (PsbO, PsbU and PsbV) associated with oxygen-evolving PSII [11]. One unusual feature was the persistence of a FLAG-ChlF/D2 aggregate after SDS-PAGE (Fig. 1a). Analysis by mass spectrometry confirmed the identity of the proteins (Supplementary Table 2) and unequivocally showed the presence of ChlF in the preparation (44 specific peptides detected covering 40% of the whole sequence). Control purifications confirmed that FLAG-tagged D1 was still able to assemble into monomeric and dimeric PSII complexes and bind PsbO and PsbV although PsbU was depleted (Fig. 1b). The absorption spectra of the two types of PSII complex were similar with a peak in the red at 672-673 nm (Fig. 1c). HPLC revealed that the pigment content of the FLAG-ChlF PSII-like complex was similar to that of typical PSII complexes (Table 1), except for the presence of low levels of Chl *f* (Fig. 1d), which was most probably responsible for the enhanced fluorescence emission of the complex at 718 nm (Fig. 1e). In contrast, FLAG-D1 complexes

gave emission bands at 685 and 694 nm expected for Chl *a*-containing PSII[12] (Fig. 1e). The emission band at 718 nm, assigned to Chl *f*, was found in the monomeric PSII core-like complex containing ChlF and, to a lesser degree, in the RC47' complex, suggesting a location for Chl *f* at the interface of CP43 and ChlF (Fig. 1f). A similar PSII-like complex was purified when FLAG-ChlF was isolated from a strain constructed in the WT Syn6803 background (Supplementary Fig. 2), except that trace amounts of a PSI/RC47' complex were present[13] (Extended Data Fig. 1). Overall these data indicate that ChlF can substitute for D1 to form modified PSII-like core complexes, which we term super-rogue PSII complexes on the basis that they contain a super-rogue version of D1 (ChlF) rather than the typical form of D1.

Testing Chl *f* production in PSII mutants

To test whether Chl *f* accumulation was dependent on the incorporation of ChlF into a PSII-like core complex, we expressed FLAG-ChlF in a series of strains unable to synthesize one or more of the PSII core subunits and so blocked at various stages of PSII assembly. In all cases Chl *f* could not be detected by HPLC, despite the expression of FLAG-ChlF and its incorporation into various PSII sub-complexes (Supplementary Table 1). These included the RC47' complex in a mutant lacking CP43 (Extended Data Fig. 2) and a novel RC43' complex, consisting of a FLAG-ChlF/D2 sub-complex attached to CP43, which accumulated in a strain unable to make CP47 (Extended Data Fig. 3). We could obtain no clear evidence for the formation of a ChlF homodimer when FLAG-ChlF was expressed in a defined PSII$^-$ strain lacking CP47, CP43, D1 and D2 (Extended Data Fig. 4 and Supplementary Fig. 5). Overall these results suggest that Chl *f* accumulation in Syn6803 requires ChlF to be assembled into a PSII-like core complex. Our results do not agree with the recent conclusion that Chl *f* can still be produced in a D2 deletion mutant of *Synechococcus* sp. PCC 7002, which was constructed by deleting the *psbD1* and *psbD2* genes[6]. Given that *psbD1* overlaps the *psbC* gene encoding CP43, such a D2 deletion strain is expected to lack CP43 and, based on analysis of a similar mutant in Syn6803,

is likely to have low levels of CP47 and D1[14]. Unlike the PSII$^-$ strain constructed here, validation of the genotype of the D2 deletion strain was not presented.

Analysis of ChlF/D1 chimeric proteins

We next tested whether it was possible to convert the D1 subunit found in oxygen-evolving PSII complexes into a Chl *f* synthase. The sequence identity between ChlF and D1 is 56%, with the main differences found at the N-terminal tail; the C-terminal tail involved in ligating the Mn_4CaO_5 oxygen-evolving cluster in PSII; the iron-lid region between transmembrane helices IV and V, which is located close to the non-heme iron and the binding pocket of the Q_B plastoquinone electron acceptor; plus two residues (QD at positions 149 and 150 in ChlF) found in the middle of the second transmembrane helix which are totally conserved in all known ChlF sequences (Fig. 2a, Supplementary Fig. 6). A series of chimeric Syn6803 mutants was constructed in which portions of ChlF were introduced into D1 and the levels of mutated D1 and Chl *f* assessed (Fig. 2). Remarkably, cells of the D1-QD mutant were able to accumulate Chl *f* to about 0.1% of Chl *a,* suggesting that the QD motif is an important determinant for Chl *f* production. The reciprocal mutation, in which the QD motif in ChlF was converted to the MG pair found in D1, blocked Chl *f* accumulation, despite assembly of FLAG-ChlF into PSII complexes (Extended Data Fig. 5). His-tagged PSII complexes isolated from the D1-QD mutant (Fig. 3a), in contrast to WT His-tagged PSII complexes (Fig. 3b), contained Chl *f* (Fig. 3c, Table 1, Supplementary Fig. 3c) and gave rise to a small but distinct fluorescence emission shoulder at 715 nm (Fig. 3d). The PSII complexes isolated from D1-QD lacked the extrinsic subunits associated with the oxygen-evolving PSII complexes and contained RC47 complexes, indicating weaker attachment of CP43 (Fig. 3a).

Identification of a possible Chl *f*-binding site

In the crystal structure of cyanobacterial PSII, the MG residues at positions 127 and 128 of D1, equivalent to Q149/D150 of ChlF, point towards a Chl *a* molecule (Chl 32 according to numbering scheme in ref. 15) at the interface of CP43 and D1 (Fig. 3e). Chl *f* is an oxidized derivative of Chl *a* in which the methyl at position 2 is converted to a formyl group (Supplementary Fig. 1). Modelling of QD into the cyanobacterial PSII crystal structure[15] suggests that these residues could H-bond to the formyl group of a Chl *f* molecule bound at this site and hence stabilize binding (Fig. 3f). The chemistry behind the conversion of Chl *a* to Chl *f* (or conceivably chlorophyllide *a* to chlorophyllide *f*) is currently unknown but the process is light-driven and has been suggested to involve a free-radical mechanism involving the activation of molecular oxygen[16]. The proposed Chl *f* binding site points to the possible involvement of reactive oxygen species (ROS) generated within the core complex following light-induced charge separation[17]. Indeed, the proposed Chl *f* binding site is located 13 Å from the redox-active pheophytin *a* electron acceptor in PSII, a site known to produce ROS[17].

Our work has clearly shown that ChlF is able to substitute for D1 to form a modified PSII complex, termed 'super-rogue' PSII complex, rather than forming a homodimer. There is no evidence for the assembly of an oxygen-evolving metal cluster or attachment of extrinsic proteins as seen in oxygen-evolving PSII, but it is possible that an alternative electron donor might be operating *in vivo*. The very low expression of the hybrid ChlF-containing PSII-like complex might explain why it has evaded detection and purification in other studies[6]. Given that ChlF/PsbA4 is a paralogue of D1, our data would also suggest that other divergent forms of D1, termed rogue D1[5], of unknown function and commonly found in nitrogen-fixing cyanobacteria, are likewise assembled into modified PSII complexes as suggested in previous work[18,19]. We have so far been unable to develop an *in vitro* system to produce Chl *f* (see methods) so the participation of other cyanobacterial factors in Chl *f* synthesis cannot yet be excluded. Chl 32, at the interface of D1 and CP43, appears to be a possible binding site for Chl *f* within the ChlF

complex and so detachment of CP43 from ChlF to release Chl *f* might be an important factor to improve the yield of Chl *f* accumulation in heterologous systems such as higher plant chloroplasts. Whether the incorporation of Chl *f* into the photosynthetic apparatus in the chloroplast would require extensive remodeling of the photosystems as observed in cyanobacteria[20] or increase susceptibility to photoinhibition due to the enhanced production of ROS in PSII through charge recombination[21] remains to be tested.

Methods

Strains and cultivation conditions

The following strains of *Synechocystis* sp. PCC 6803 (hereafter Syn6803) were used in this work: the glucose-tolerant wild-type strain GT-P (hereafter denoted WT)[22]; *psbA* triple-deletion strains (ΔPsbA)[23] and 4E-3H6 (hereafter His-CP47/ΔD1)[24]; the D1 and CP43 deletion strain, ΔD1/ΔCP43; the D1 and CP47 deletion strain, ΔD1/ΔCP47; the PSII-less strain, ΔD1/ΔD2/ΔCP43/ΔCP43, denoted here PSII$^-$. To create the ΔD1/ΔCP43 strain, the *psbD1C* operon was removed from the ΔD1 strain[23] by replacing the open reading frame of *psbD1* (encoding D2) and *psbC* (encoding CP43) with a kan^R-SacB selection cassette[26]. Once the fully segregated *psbD1C* deletion mutant was obtained by selecting on kanamycin and confirmed by PCR, a new DNA template containing only the *psbD1C* flanking regions, separated by an EcoRV restriction site (GATATC), was used to transform the strain, in order to remove the previously introduced selection cassette. Negative selection was performed by supplementing the BG11 medium with 5% (w/v) sucrose in the absence of kanamycin. The ΔD1/ΔCP47 deletion strain was also derived from the ΔD1 deletion strain[23] by replacing the *psbB* gene (encoding CP47) with a gentamycin-resistance cassette. The PSII-less strain was derived from the ΔD1/ΔCP43 deletion strain by sequential deletion of *psbB* as described above

and then *psbD2*, encoding D2, by an erythromycin-resistance cassette using standard transformation protocols[27].

All Syn6803 strains were cultivated on BG-11 agar plates (BG-11 basic mineral medium, supplemented with 0.3 % (w/v) sodium thiosulfate, 10 mM N-tris[hydroxymethyl]methyl-2-aminoethanesulfonic acid (TES-KOH) pH = 8.2 (ref 27) and 1.5 % (w/v) agar; 5 mM glucose was included when necessary[27]. Liquid BG-11 cultures (buffered with 5 mM TES-KOH pH = 8.2 and containing 5 mM glucose and appropriate antibiotics when necessary) were grown in sterile, air-filter capped tissue culture flasks (Thermo Fisher Scientific, USA) on an orbital shaker incubator (Innova 2100 shaker, New Brunswick Scientific, UK) at 100 rpm. The strains were incubated in a temperature-controlled room set at 29-30 °C and illuminated with fluorescent white light of intensity 10-30 µmol photons m^{-2} s^{-1}. Antibiotics for selective growth were used at the following concentrations: kanamycin (10-50 µg ml^{-1}), chloramphenicol (30 µg ml^{-1}), gentamycin (5 µg ml^{-1}).

Construction of *chlF* and *psbA2* mutants

The pPD FLAG vector was used to express ChlF from the *psbA2* locus in Syn6803, either in its native form or with an N-terminal 3xFLAG-tag[28]. The *chlF*/*psbA4* gene from *Chroococcidiopsis thermalis* PCC 7203 was amplified using a Phusion® High Fidelity PCR Master Mix (New England Biolabs, USA) using genomic DNA as template and several different sets of primers (Supplementary Table 3). Plasmids were constructed by In-fusion cloning[29]. The resulting plasmids, designated pPD-ChlF and pPD FLAG-ChlF, were used to transform WT-G, ΔD1, ΔD1/ΔCP43, ΔD1/ΔCP47, and His-CP47/ΔD1 strains of Syn6803 to yield ChlF and FLAG-ChlF mutants.

The chimeric D1/ChlF mutants were generated by in-fusion cloning in which portions of *chlF* replaced *psbA2* (encoding D1) present in plasmid pRD1031Km^R (ref 25), except for the N-

terminal tail mutant, which used the pPD FLAG vector[28]. The following mutants were made (Fig. 2B, Supplementary Fig. 6): D1 with the N-terminus of FLAG-ChlF (denoted N-terminal tail mutant), D1 with the C-terminal region of ChlF (C-terminal tail mutant), D1 in which the iron lid region was replaced with analogous ChlF sequences (Iron lid mutant), D1 mutant in which the MG pair at positions 127 and 128 was mutated to QD to mimic ChlF (D1-QD mutant), D1 with region from QD to the C-terminal tail of ChlF (QD_C-terminal tail mutant). The primers used to generate the mutants are listed in Supplementary Table 3. All plasmid constructs were used to transform the Syn6803 His-CP47/ΔD1 strain.

Genotyping

The genotype of *Synechocystis* mutants was verified by PCR amplification from genomic DNA using appropriate PCR primers. The genomic DNA of *Synechocystis* mutants was extracted using *Quick*-DNA™ Fungal/Bacterial Miniprep Kit (Zymo Research, USA), followed by PCR amplification. Once segregation was confirmed, the DNA fragments were gel purified and sequenced to confirm the correct sequence.

Spectroscopic methods

The absorption spectra of cells were measured using a UV-3000 spectrophotometer (Shimadzu, Japan). The 77 K Chl fluorescence spectra were recorded using an Aminco Bowman Series 2 luminescence spectrometer (Spectronic Unicam, USA). The frozen sample was excited at 435 nm and the fluorescence emission was recorded between 600 and 800 nm.

Preparation of cellular membranes

Cells in the exponential growth phase (~5 x 10^9 cells per ml) were pelleted, washed, and re-suspended in buffer B (25 mM MES/NaOH, pH 6.5, 10 mM $CaCl_2$, 10 mM $MgCl_2$, 25 % (v/v) glycerol). Cells were broken mechanically in a Mini-Beadbeater (BioSpec, USA) using glass beads (100-200 µm diameter) as described in ref 30. The membrane and soluble protein

fractions were separated by centrifugation (Sigma 3K30, Germany; 65,000 *x g*, 20 min). The pelleted membranes were washed and re-suspended in buffer B.

Isolation of protein complexes

For the purification of FLAG-tagged protein complexes under native conditions, cells from 4 L cultures (OD at 750 nm of about 0.5) were centrifuged, washed and broken as described above. The cellular membranes containing 1 mg/mL Chl were solubilized for 1 h with 1 % (w/v) n-dodecyl-β-D-maltoside (DDM) at 10 °C, then centrifuged for 20 min at 65,000 x g to remove insolubilized membrane particles. The supernatant was loaded onto an anti-FLAG M2 affinity gel chromatography column (Sigma-Aldrich, Germany). Proteins bound to the column were washed 6 times with 2.5 resin volumes of buffer B containing 0.04 % (w/v) DDM. FLAG-tagged protein complexes were eluted with buffer B containing 0.04 % (w/v) DDM and 150 µg/mL 3xFLAG peptide (Sigma-Aldrich, Germany).

His-tagged CP47 preparations were isolated using the immobilized Ni-affinity chromatography protocol described in ref 31 designed to purify highly active oxygen-evolving PSII complexes.

Protein electrophoresis and immunoblotting

For native electrophoresis, solubilized membrane proteins or isolated complexes were separated on 4 to 12 % (w/v) polyacrylamide clear native (CN)[32] gels in the 1st dimension. The protein gels were scanned and the Chl fluorescence image was taken by a LAS 4000 camera (Fuji, Japan). The individual components of the protein complexes were resolved by incubating the gel strip from the first dimension in 62.5 mM Tris/HCl pH 6.8 containing 2 % (w/v) SDS and 1 % (w/v) dithiothreitol for 30 min at room temperature, and by subsequent separation in the 2nd dimension by SDS-electrophoresis in a denaturing 12 to 20 % (w/v) polyacrylamide gel containing 7 M urea[33]. For standard single dimension SDS-PAGE, membrane suspensions were solubilized at room temperature for 20 min after adding 1/5 volume of 10 % (w/v) SDS and 5 %

(w/v) dithiothreitol. Proteins were stained by Coomassie Brilliant Blue (CBB) (Sigma-Aldrich, Germany) or, alternatively, in the case of subsequent immunoblotting, by SYPRO Orange (Sigma-Aldrich, Germany). Protein from the SYPRO-stained gel was transferred onto a PVDF membrane that was subsequently incubated with a specific primary antibody and then with secondary antibody conjugated with horseradish peroxidase (Sigma-Aldrich, Germany). The blot was visualized using the chemiluminescent substrate Immobilon Crescendo (Millipore, USA) and the signal was recorded by a LAS-4000 imaging system (Fuji, Japan). The following primary antibodies were used: anti-FLAG (Abgent, cat. no. AP1013A); anti-D1 (N-terminal specific)[34]; anti-D2[35], anti-CP43 (Agrisera, cat. no. AS11 1787), anti-CP47 (Agrisera, cat. no. AS04 038), anti-PsbF[36], anti-PsbH[37], anti-PsaD[38], anti-PsbI[39]; anti-Ycf48[40] and anti-Psb27[41]

Enzymatic digestion and protein identification by mass spectrometry

For identification of proteins by LC-MS/MS, the CBB-stained protein bands were cut from the gel and digested by trypsin. Resulting peptides were extracted, purified with ZipTip C18 pipette tips (Millipore, USA) and analyzed using a NanoAcquity UPLC (Waters, USA) on-line coupled to an ESI Q-ToF Premier mass spectrometer (Waters, USA) as described in detail in ref 42. The whole preparation was analyzed using the same procedure after previous protein precipitation using 2D clean-up kit (GE Healthcare, USA) according to the manufacturer's instructions.

Spectrophotometric determination of Chl contents

The Chl content per cell was determined spectrophotometrically using a UV-3000 spectrophotometer (Shimadzu, Japan) after extraction of pigments with 100% methanol as described in ref 43.

Pigment quantification

In order to determine molar stoichiometries of pigment cofactors in purified protein complexes, we divided 40 µL of the analysed sample into halves. The first half of the sample was used to

quantify heme-*b* content, the second half for the measurement of Chl, β-carotene and pheophytin. For the heme measurement, 20 µL of the sample was mixed with 200 µL of acetone/0.2% HCl and the supernatant collected after centrifugation for 5 min at 12,000 rpm. 200 µL of the extract was immediately injected onto an HPLC machine (Agilent-1200) and separated using a reverse phase column (C18, 4 µm particle size, 3.9 x 75 mm, Waters) with 0.1% (v/v) trifluoroacetic acid in water and 0.1% trifluoroacetic acid/acetonitrile, as solvents A and B, respectively. Heme-*b* was eluted with a linear gradient of solvent B (30–100% in 12 min) followed by 100% of B at a flow rate of 0.8 ml/min at 40 °C and detected by a diode array detector at 400 nm. To determine the content of Chl, β-carotene and pheophytin *a*, 20 µL of the sample was mixed with 200 µL of methanol and centrifuged for 5 min at 12,000 rpm. 200 µL of the collected supernatant was separated on a reverse phase column (Zorbax Eclipse C18, 5 µm particle size, 3.9 × 150 mm; Agilent) with 35% (v/v) methanol and 15% (v/v) acetonitrile in 0.25 M pyridine (solvent A) and 20% (v/v) methanol, 20% (v/v) acetone, 60% (v/v) acetonitrile as solvent B. Pigments were eluted with a linear gradient of solvent B (30–95% (v/v) in 25 min) in solvent A followed by 95% of solvent B in solvent A at a flow rate of 0.8 ml min^{-1} at 40 °C. Chl *a* and β-carotene were detected by a diode array detector at 665 nm and 450 nm, Chl *f* was detected at 710 nm. Pheophytin *a* was detected by a fluorescence detector using 415 nm and 680 nm as excitation and emission wavelengths, respectively. The obtained peaks were integrated and the molar stoichiometries calculated using molar coefficients. For Chl *a*, β-carotene and pheophytin *a,* these coefficients were determined using calibration curves prepared with authentic standards dissolved in methanol. The Chl *f* to Chl *a* molar ratio was estimated by integrating the Q_y maxima (710 and 665 nm)[44]. However, due to a problem with the solubility of hemin, we did not obtain an accurate molar extinction coefficient for heme *b.* We therefore used a highly pure preparation of monomeric PSII from *Thermosynechococcus elongatus* (obtained from Marc Nowaczyk, University of Bochum) for heme *b* calibration based on the known ratio of heme *b* and pigment in PSII (1 heme *b* per 2 pheophytin *a* and 35 Chl *a*).

Testing the synthesis of Chl *f in vitro*

Attempts to synthesise Chl *f in vitro* were made using the FLAG-ChlF preparation isolated from the FLAG-ChlF/ΔD1 strain. The preparation was diluted to a final concentration of 1 µg/ml of Chl *a* in different buffers containing various additives and incubated in the dark or in the light (40 µmol photons m^{-2} s^{-1}) for one or two hours at 27 ºC. We used either buffer B or 25 mM HEPES/NaOH, pH 7.5, 10 mM $CaCl_2$, 10 mM $MgCl_2$, both additionally containing 0.05% (w/v) DDM and all possible combinations of the following substances: Chl *a* in methanol (final concentration 1 µg/ml) as a substrate, 1 mM diphenyl carbazide or 1 mM $MnCl_2$ or 1 mM hydroxylamine as an electron donor, and 1 mM potassium ferricyanide as an electron acceptor. The amount of Chl *f* was assessed using HPLC before and after incubation and was compared with the identically treated control mixtures lacking the FLAG-ChlF preparation. In all cases no measurable increase in Chl *f* concentration was observed.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon request.

Statistics and reproducibility

Figure 1: Three independent isolations of Flag-D1 and Flag-ChlF were performed and their analyses (2D gel analysis in Fig. 1a and b, absorption spectra in Fig. 1c, pigment analyses in Fig. 1d, 77K fluorescence spectra in Fig. 1e and absorption spectra of gel bands in Fig. 1f) gave very similar and reproducible results for the two types of sample.

Figure 2c: The stained SDS-PAGE gel was done once; the membranes were immunoblotted twice and gave reproducible results.

Figure 3a-d: Five independent isolations of His-CP47/D1-QD complexes and three isolations of His-CP47 complexes were performed and their analyses (2D gel analysis in Fig. 3a and b, pigment analyses in Fig. 3c, 77 K fluorescence spectra in Fig. 3d) gave very similar and reproducible results for each type of sample.

Extended Data Fig. 1a: Analyses were performed twice with similar results.

Extended Data Fig. 2b-c: The PCR analysis in panel b was performed twice with similar results and the gel analysis in panel c was performed twice with similar results.

Extended Data Fig. 3b-c: The PCR analysis in panel b was performed twice with similar results and the gel analysis in panel c was performed twice with similar results.

Extended Data Fig. 4a-b: Analyses were performed twice with similar results.

Extended Data Fig. 5: Analysis was performed twice with similar results.

Supplementary Figure 2b: The PCR analysis was repeated independently three times with similar results.

Supplementary Figure 3a-c: The analyses and spectra measurements were performed twice for *Chroococcidiopsis thermalis* (Supplementary Fig. 3a) with similar results, three times for Flag-ChlF with similar results and five times for His-CP47/D1-QD with similar results.

Supplementary Figure 4a-b: Analyses were performed twice with similar results.

Supplementary Figure 5: The PCR analysis was performed twice with similar results

Supplementary Table 1: Numbers represent mean ± SD for independent preparations. For WT and ΔD1, n = 2; for His-CP47/D1-QD, n = 5; for remaining strains and preps, n = 3.

References

1. Miyashita, H. *et al.* Chlorophyll d as a major pigment Chlorophyll d as a major pigment. *Nature* **383**, 402–402 (1996).

2. Chen, M., Li, Y., Birch, D. & Willows, R. D. A cyanobacterium that contains chlorophyll f - a red-absorbing photopigment. *FEBS Lett.* **586**, 3249–3254 (2012).

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

4. Ho, M.-Y., Shen, G., Canniffe, D. P., Zhao, C. & Bryant, D. A. Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II. *Science (80-.).* **353**, aaf9178–aaf9178 (2016).

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

6. Shen, G. *et al.* Characterization of chlorophyll f synthase heterologously produced in *Synechococcus* sp. PCC 7002. *Photosynth. Res.* **140**, 77–92 (2019).

7. Tilman, D., Balzer, C., Hill, J. & Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci.* **108**, 20260–20264 (2011).

8. Cardona, T., Shao, S. & Nixon, P. J. Enhancing photosynthesis in plants: the light reactions. *Essays Biochem.* **62**, 85–94 (2018).

9. Long, S. P., Marshall-Colon, A. & Zhu, X.-G. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161**, 56–66 (2015).

10. Nowaczyk, M. M. *et al.* Psb27, a cyanobacterial lipoprotein, is involved in the repair cycle of Photosystem II. *Plant Cell* **18**, 3121–3131 (2006).

11. Barber, J. Photosystem II: the water splitting enzyme of photosynthesis and the origin of

oxygen in our atmosphere. *Q. Rev. Biophys.* **49**, e14 (2016).

12. Shen, G., Boussiba, S. & Vermaas, W. F. J. *Synechocystis* sp PCC 6803 strains lacking Photosystem I and phycobilisome function. *Plant Cell* **5**, 1853–1863 (2007).

13. Bečková, M. *et al.* Association of Psb28 and Psb27 proteins with PSII-PSI supercomplexes upon exposure of *Synechocystis* sp. PCC 6803 to high light. *Mol. Plant* **10**, 62–72 (2017).

14. Komenda, J. *et al.* Accumulation of the D2 protein is a key regulatory step for assembly of the photosystem II reaction center complex in *Synechocystis* PCC 6803. *J. Biol. Chem.* **47**, 48620–48629 (2004).

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

16. Garg, H., Loughlin, P. C., Willows, R. D. & Chen, M. The $C2^1$ -formyl group in chlorophyll f originates from molecular oxygen. *J. Biol. Chem.* **292**, 19279–19289 (2017).

17. Vass, I. Molecular mechanisms of photodamage in the Photosystem II complex. *Biochim. Biophys. Acta* **1817**, 209–217 (2012).

18. Masuda, T. *et al.* Diel regulation of photosynthetic activity in the oceanic unicellular diazotrophic cyanobacterium *Crocosphaera watsonii* WH8501. *Environ. Microbiol.* **20**, 546–560 (2018).

19. Wegener, K. M., Nagarajan, A. & Pakrasi, H. B. An atypical *psbA* gene encodes a sentinel D1 protein to form a physiologically relevant inactive photosystem II complex in cyanobacteria. *J. Biol. Chem.* **290**, 3764–3774 (2015).

20. Gan, F. *et al.* Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light. *Science* **345**, 1312–1317 (2014).

21. Nürnberg, D. J. *et al*. Photochemistry beyond the red-limit in chlorophyll f-containing systems. Science **360**, 1210-1213 (2018).

22. M. Tichý, M. Bečková, J. Kopečná, J. Noda, R. Sobotka, J. Komenda, Strain of *Synechocystis* PCC 6803 with aberrant assembly of photosystem II contains tandem duplication of a large chromosomal region. *Front. Plant Sci.* **7,** 648 (2016).

23. Forsman, J. A., Winter, R. T. & Eaton-Rye, J. J. An improved system for the targeted mutagenesis of the *psbA2* gene in *Synechocystis* sp. PCC 6803: mutation of D1-Glu244 to His impairs electron transfer between Q_A and Q_B of photosystem II. *New Zeal. J. Bot.* **57**, 125–136 (2019).

24. Debus, R. J. *et al.* Does histidine 332 of the D1 polypeptide ligate the manganese cluster in photosystem II? An electron spin echo envelope modulation study. *Biochemistry* **40**, 3690–3699 (2001).

25. Debus, R. J., Barry, B. A., Sithole, I., Babcock, G. T. & Mcintosh, L. Directed mutagenesis indicates that the donor to $P680^+$ in photosystem II is tyrosine-161 of the D1 polypeptide. *Biochemistry* **27**, 9071–9074 (1988).

26. Ried, J. L. & Collmer, A. An nptI-sacB-sacR cartridge for constructing directed, unmarked mutations in Gram-negative bacteria by marker exchange-eviction mutagenesis. *Gene* **57**, 239–246 (1987).

27. Williams, J. G. K. Construction of specific mutations in photosystem II photosynthetic reaction center by genetic engineering methods in *Synechocystis* 6803. in *Methods in Enzymology* **167**, 766–778 (1988).

28. Hollingshead, S. *et al.* Conserved chloroplast open-reading frame ycf54 is required for activity of the magnesium protoporphyrin monomethylester oxidative cyclase in *Synechocystis* PCC 6803. *J. Biol. Chem.* **287**, 27823–27833 (2012).

29. Zhu, B., Cai, G., Hall, E. O. & Freeman, G. J. In-Fusion™ assembly: Seamless engineering of multidomain fusion proteins, modular vectors, and mutations. *Biotechniques* **43**, 354–359 (2007).

30. Komenda, J. & Barber, J. Comparison of *psbO* and *psbH* deletion mutants of *Synechocystis* PCC 6803 indicates that degradation of D1 protein is regulated by the Q_B site and dependent on protein synthesis. *Biochemistry* **34**, 9625–9631 (1995).

31. Suzuki, H. *et al.* Functional roles of D2-Lys317 and the interacting chloride ion in the water oxidation reaction of photosystem II as revealed by fourier transform infrared analysis. *Biochemistry* **52**, 4748–4757 (2013).

32. Wittig, I., Karas, M. & Schägger, H. High resolution clear native electrophoresis for in-gel functional assays and fluorescence studies of membrane protein complexes. *Mol. Cell. Proteomics* **6**, 1215–1225 (2007).

33. Dobáková, M., Sobotka, R., Tichý, M. & Komenda, J. Psb28 protein is involved in the biogenesis of the photosystem II inner antenna CP47 (PsbB) in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol.* **149**, 1076–1086 (2009).

34. Komenda, J., Tichý, M. & Eichacker, L. A. The PsbH protein is associated with the inner antenna CP47 and facilitates D1 processing and incorporation into PSII in the cyanobacterium *Synechocystis* PCC 6803. *Plant Cell Physiol.* **46**, 1477–1483 (2005).

35. Rögner, M., Nixon, P. J. & Diner, B. A. Purification and characterization of photosystem I and photosystem II core complexes from wild-type and phycocyanin-deficient strains of the cyanobacterium *Synechocystis* PCC 6803. *J. Biol. Chem.* **265**, 6189–96 (1990).

36. Komenda, J. *et al.* The cyanobacterial homologue of HCF136/YCF48 is a component of an early photosystem II assembly complex and is important for both the efficient assembly and repair of photosystem II in *Synechocystis* sp. PCC 6803. *J. Biol. Chem.*

283, 22390–22399 (2008).

37. Bumba, L., Prasil, O. & Vacha, F. Antenna ring around trimeric Photosystem I in chlorophyll b containing cyanobacterium *Prochlorothrix hollandica*. *Biochim. Biophys. Acta - Bioenerg.* **1708**, 1–5 (2005).

38. Li, N. *et al.* PsaD is required for the stable binding of PsaC to the photosystem I core protein of *Synechococcus* sp. PCC 6301. *Biochemistry* **30**, 7863–7872 (1991).

39. Dobáková, M., Tichý, M. & Komenda, J. Role of the PsbI protein in photosystem II assembly and repair in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol.* **145**, 1681–1691 (2007).

40. Yu, J. *et al.* Ycf48 involved in the biogenesis of the oxygen-evolving photosystem II complex is a seven-bladed beta-propeller protein. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E7824–E7833 (2018).

41. Komenda, J. *et al.* The Psb27 assembly factor binds to the CP43 complex of photosystem II in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol.* **158**, 476–486 (2012).

42. Bučinská, L. *et al.* The ribosome-bound protein Pam68 promotes insertion of chlorophyll into the CP47 subunit of photosystem II. *Plant Physiol.* **176**, 2931–2942 (2018).

43. Ritchie, R. J. Consistent sets of spectrophotometric chlorophyll equations for acetone, methanol and ethanol solvents. *Photosynth. Res.* **89**, 27–41 (2006).

44. Li, Y., Scales, N., Blankenship, R. E., Willows, R. D. & Chen, M. Extinction coefficient for red-shifted chlorophylls: Chlorophyll d and chlorophyll f. *Biochimica et Biophysica Acta - Bioenergetics* **1817**, 1292–1298 (2012).

All correspondence and requests for materials should be addressed to Peter J. Nixon.

Acknowledgments

We thank Ben Nwaobi for excellent laboratory management and Richard Debus and Julian Eaton-Rye for providing plasmids and strains.

Funding

Financial support was provided by the BBSRC (BB/P00931X/1 to J.W.M. and P.J.N.), the Grant Agency of the Czech Republic (project no. 19-29225X to M.B., R.S. and J.K.) and the Ministry of Education, Youth and Sports of the Czech Republic (National Program of Sustainability I, ID: LO1416 to M.B., R.S., and J.K.). J.P.T received a PhD scholarship from the Indonesia Endowment Fund for Education (LPDP).

Author contributions

P.J.N., J.K. and J.W.M. conceived the research, J.P.T., S.S and J.K. prepared the figures and P.J.N and J.K. wrote the manuscript with contributions from all the other authors. P.J.N. and J.K. coordinated the activities. J.P.T. constructed the mutants. J.P.T., M.B., S.S., J.Y., Z.Z., R.S. performed the biochemical analyses. All authors approved the final version of the manuscript.

Competing interests: Authors declare no competing interests.

Figure Legends

Fig. 1. Isolation and characterization of ChlF complexes. (a) FLAG-ChlF and (b) FLAG-D1 complexes were immunoaffinity purified and analyzed by 2D gel electrophoresis. Pigmented complexes, separated by clear-native (CN) gel electrophoresis in the first dimension (1D color), were visualized by fluorescence (1D fluor) then denatured, separated by SDS-PAGE in the second dimension and protein subunits detected by Coomassie Brilliant Blue staining (CBB). Proteins were identified by immunoblotting and mass spectrometry (Supplementary Fig. 4, Supplementary Table 2). (c) Room-temperature spectra of FLAG-tagged complexes with spectra offset for clarity. (d) Detection of Chl *a* and Chl *f* by HPLC and absorbance at 710 nm measured using a diode array detector (DAD). 77 K fluorescence emission spectra (440 nm excitation) of (e) FLAG-ChlF and FLAG-D1 complexes and (f) FLAG-ChlF complexes and sub-complexes separated by CN gel electrophoresis. Abbreviations: F.ChlF, FLAG-tagged ChlF; F.D1, FLAG-tagged D1; PSII(1) and PSII(2), monomeric and dimeric PSII complexes, respectively; RC47, PSII(1) lacking CP43; PSII(1)' and RC47' are PSII-like complexes containing ChlF rather than D1; FP, free pigment.

Fig.2. Expression of chimeric ChlF/D1 proteins in Syn6803. (a) Folding model of D1 protein indicating regions replaced by ChlF sequences. (b) Cartoon showing regions of D1 replaced by ChlF sequences in each strain and level of Chl *f* detected in cells (% Chl *a*); nd, not detectable. (c) Immunoblot using N-terminal D1-specific antibody (α D1) of membrane proteins isolated from mutants and the corresponding Coomassie Brilliant Blue stained gel (Coomassie).

Fig.3. Characterization of PSII complexes isolated from the D1-QD mutant. His-tagged CP47 complexes were isolated from (a) His-CP47/D1-QD and (b) WT control strain (His-CP47)

and analyzed by 2D gel electrophoresis as in Fig. 1. (c) Detection of Chl *f* in D1-QD complexes by HPLC and absorbance at 710 nm using a diode array detector (DAD). (d) 77 K fluorescence emission spectra (440 nm excitation) of the monomeric PSII complexes separated by CN electrophoresis in the His-CP47/D1-QD sample in panel a (PSII(1)/D1-QD) and His-CP47 sample in panel b (PSII(1)). (e) Determined location of residues D1-M127 and D1-Gly128 and Chl 32 in the structure of PSII (PDB:3WU2). (f) Possible interaction between the formyl group of Chl *f* and D1-QD residues after replacing the MG pair by QD and Chl *a* by Chl *f*.

Table 1. Pigment analysis of affinity-purified FLAG-ChlF, FLAG-D1 and His-CP47/D1-QD preparations. The ratio of all pigments is normalized to the single heme *b* found in PSII.

	Stoichiometry of pigment cofactors				
Preparation	Chl *a*	Chl *f*[a]	β-carotene	pheophytin	Heme *b*
FLAG-ChlF/ΔD1	33.7 ± 0.3	0.33 ± 0.02.	7.1 ± 0.3	2.4 ± 0.2	1
FLAG-D1	35.3 ± 0.1	0	8.7 ± 0.1	2.3 ± 0.2	1
His-CP47/D1-QD	33.0 ± 1.0	0.08 ± 0.02	7.3 ± 0.7	2.1 ± 0.2	1
His-CP47	39.0 ± 1.0	0	9.7 ± 1	2.6 ± 0.25	1

[a]An approximate value calculated from the ratio between the absorbance of Q_y peaks of Chl *a* and Chl *f* recorded at 665 and 710 nm, respectively. Numbers represent mean ± SD for independent preparations of His-CP47/D1-QD (n =5) and the FLAG-tagged samples (n=3). For the His-CP47 control, the numbers represent the mean ± SD of 3 technical replicates of a single preparation.