The FEBS Journal

FEBS PRESS
science publishing by scientists

STRUCTURAL SNAPSHOT

Structural features of the diatom photosystem II–light-harvesting antenna complex

Wenda Wang[1], Songhao Zhao[1,2], Xiong Pi[3], Tingyun Kuang[1], Sen-Fang Sui[3] and Jian-Ren Shen[1,4]

1 Photosynthesis Research Center, Key Laboratory of Photobiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China
2 University of Chinese Academy of Sciences, Beijing, China
3 State Key Laboratory of Membrane Biology, Beijing Advanced Innovation Center for Structural Biology, School of Life Sciences, Tsinghua University, Beijing, China
4 Research Institute for Interdisciplinary Science and Graduate School of Natural Science and Technology, Okayama University, Japan

Keywords
diatom; energy transfer; fucoxanthin; light harvesting; photosystem

Correspondence
J. -R. Shen, Photosynthesis Research Center, Key Laboratory of Photobiology, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China and Research Institute for Interdisciplinary Science, Okayama University, Okayama, Japan
Tel: +86-10-62836231
E-mail: jrshen@ibcas.ac.cn

(Received 6 November 2019, revised 5 December 2019, accepted 17 December 2019)

doi:10.1111/febs.15183

In photosynthesis, light energy is captured by pigments bound to light-harvesting antenna proteins (LHC) that then transfer the energy to the photosystem (PS) cores to initiate photochemical reactions. The LHC proteins surround the PS cores to form PS-LHC supercomplexes. In order to adapt to a wide range of light environments, photosynthetic organisms have developed a large variety of pigments and antenna proteins to utilize the light energy efficiently under different environments. Diatoms are a group of important eukaryotic algae and possess fucoxanthin (Fx) chlorophyll *a*/*c* proteins (FCP) as antenna which have exceptional capabilities of harvesting blue-green light under water and dissipate excess energy under strong light conditions. We have solved the structure of a PSII–FCPII supercomplex from a centric diatom *Chaetoceros gracilis* by cryo-electron microscopy, and also the structure of an isolated FCP dimer from a pennate diatom *Phaeodactylum tricornutum* by X-ray crystallography at a high resolution. These results revealed the oligomerization states of FCPs distinctly different from those of LHCII found in the green lineage organisms, the detailed binding patterns of Chl *c* and Fxs, a huge pigment network, and extensive protein–protein, pigment–protein, and pigment–pigment interactions within the PSII–FCPII supercomplex. These results therefore provide a solid structural basis for examining the detailed mechanisms of the highly efficient energy transfer and quenching processes in diatoms.

Introduction

Photosynthetic light-energy conversion is achieved by two multi-subunit membrane protein complexes called photosystem I (PSI) and PSII in oxygenic organisms, and by reaction center complexes in anoxygenic photosynthetic bacteria and heliobacteria. Among the two PS, PSII catalyzes light-induced water splitting, leading to the generation of molecular oxygen and conversion of light energy to chemical energy [1–3]. To cope with the various light environments where different photosynthetic organisms inhabit, different types and numbers of light-harvesting antenna proteins are attached to the PSII core to enhance the light-harvesting and

Abbreviations
Bcr, β-carotene; Chl, chlorophyll; cryo-EM, cryo-electron microscopy; Ddx, diadinoxanthin; DGDG, digalactosyl diacylglycerol; FCP, fucoxanthin chlorophyll *a/c* protein; Fx, fucoxanthin; LHC, light-harvesting complex; MGDG, monogalactosyl diacylglycerol; NPQ, nonphotochemical quenching; PG, phosphatidylglycerol; PS, photosystem; SQDG, sulfoquinovosyl diacylglycerol.

energy-quenching capacities [2,4–10]. The pigments involved in energy harvesting are bacteriochlorophylls in anoxygenic photosynthetic bacteria [11–13], and mainly chlorophylls (Chl) *a*/*b* and various carotenoids in oxygenic photosynthetic organisms [8,10].

Diatoms are one of the important groups of oxygenic photosynthetic organisms, because they contribute around 20% of the global or 40% of the aquatic net primary production and hence play an important role in global carbon cycling [14]. Diatoms possess a huge family of light-harvesting antennae known as fucoxanthin (Fx) Chl *a*/*c*-binding proteins (FCPs), which surround the cores of both PSI and PSII to form PS-fucoxanthin (Fx) chlorophyll *a*/*c* protein (FCP) supercomplexes, where FCPs serve to harvest and transfer the solar energy to the PS cores [8,15,16]. Chl *c* and Fx associated with FCPs have unique light absorption properties different from those of Chl *b*, β-carotene (Bcr), and lutein found in the green lineage photosynthetic organisms. Chl *c* does not possess the phytol tail seen in Chl *a*/*b*, but has a conjugated system at its protoporphyrin ring that is prolonged to its propionic acid tail, which greatly enhances its absorption at the blue-green light region (460–510 nm) [17]. On the other hand, Fx has a shorter conjugated system different from that seen in luteins and Bcr, which confers FCPs the ability to harvest green light (510–560 nm) [8,15,16]. These specific pigments are important for the light harvesting of diatoms and other related algae, because more blue-green photons are available under water. In addition, a diatoxanthin–diadinoxanthin (Ddx) cycle is found in the FCP antennae which functions in quenching the excess energy under strong light conditions [7,8]. The de-epoxidation and epoxidation enzymes functioning in this cycle may be more sensitive to pH changes in the lumen than that of the enzymes involved in the zeaxanthin–violaxanthin cycle of green algae and higher plants, conferring diatoms with a higher ability of non-photochemical quenching (NPQ). This is another important feature of FCPs, since diatoms experience a more dynamic, fluctuating light environment in water due to the continuous flow and mixing of water layers, which would require a more robust system to protect diatoms from photodamage. The outstanding ability of FCPs in absorbing blue-green light and in dissipating excess energy may constitute one of the major reasons for diatoms to survive successfully in the aquatic environment [7,8,10,11].

The FCPII proteins associated with the PSII core of diatoms consist of products encoded by the *lhcf* gene family as well as some subunits encoded by *lhcx* genes which are involved in photoprotection [15,16]. All of these proteins share some sequence identities with the light-harvesting complexes (LHCs) from green algae and higher plants; however, significant differences are found in some regions of their sequences as well as their structures (see below).

The main state of FCPIIs around PSII core is previously suggested as trimer based on the LHCII trimer structure of green lineage organisms [15,18]. However, we isolated a homodimeric FCP from a pennate diatom *Phaeodactylum tricornutum* and solved its crystal structure at a high resolution [19]. This dimer is encoded by an *lhcf4* gene and may be one of the important forms of FCP existing around the PSII core. Additionally, FCPs can form some larger oligomers consisting of at least five subunits [15,16]. The large variability in the composition and oligomeric states of FCP allows a much more flexible light-harvesting antenna system in the peripheral region of the PSII core in diatoms and related algae, giving rise to a much more dynamic nature of the antenna system in these algae.

To understand the oligomeric state, organization of protein subunits, and the exact composition and distribution of pigments and other cofactors in the PSII core and its complex with antenna proteins, several PSII structures from cyanobacteria, red alga, green alga, and higher plants have been solved by cryo-electron microscopy (cryo-EM) and X-ray crystallography [1–5,20]. In particular, single-particle cryo-EM analysis has provided a powerful tool to unravel the organization of several PSII–LHC and PSI–LHC supercomplexes from different lineages of organisms from cyanobacteria to green algae and higher plants [2,4,5,21–24]. The structure of PSII–FCPII from diatoms, however, was solved only recently [25,26], and a high-resolution crystal structure of an isolated FCP dimer was also reported recently [19]. These studies revealed the oligomeric states and organization of FCPII subunits around the PSII core, the distribution and detailed binding environments of Chls *c* and Fxs, and a huge pigment–protein network within the PSII–FCPII supercomplex, which enable us to examine the excitation energy transfer and quenching pathways in the diatom PSII–FCPII in a greater detail. We discuss the structural organization of PSII–FCPII and their functional implications in the following.

Overall structure of the PSII–FCPII supercomplex

According to the symmetry of their siliceous cell walls, diatoms are divided into two groups, namely pennate and centric groups. Genome and peptide analyses show that the pennate and centric diatoms have different composition of Lhcf subunits in thylakoid membranes [15,16,19]. In the centric diatom *Chaetoceros gracilis*, FCPII exists mainly in three types designated FCP-A, FCP-B, and FCP-C [27]. The structure of a PSII–FCPII supercomplex with FCP-A as the dominant antenna was analyzed up to 3.0 Å resolution by cryo-EM [25,26]. Surprisingly, the structure shows that the major FCP exists in a tetramer (Fig. 1A,B), and there are two FCP tetramers associated with the PSII core; one is directly associated with the core at the CP47 side and designated strongly associated tetramer [S-tetramer (ST)], whereas the other one is associated with the PSII core indirectly at the CP43 side through two FCP monomers, FCP-D and FCP-E, and hence designated moderately associated tetramer [M-tetramer (MT)]. In addition, three FCP monomers are associated with each PSII core, among which FCP-F is associated at the periphery region of MT, and FCP-D and FCP-E are associated with the core directly at positions resembling that of CP29 and CP26, respectively, in PSII–LHCII supercomplexes of the green lineage organisms [2,4,5]. Thus, a total of 11 FCP subunits are associated with each PSII core.

The whole PSII–FCPII supercomplex is organized in a dimer, within which the structure of PSII core is largely similar to that from cyanobacteria to higher plants [1–5,20,28]. However, there are some notable features of the diatom PSII core different from that of cyanobacterial and higher plant PSII. The diatom PSII core is composed of 19 membrane-spanning subunits and five extrinsic proteins. Among the membrane-spanning subunits, PsbG and Psb34 are newly found and located in a marginal region of the core that is interacting with the FCP ST directly (Fig. 1A). The identity of these two subunits is not clear, but a subunit in the same position as Psb34 is also found in a red algal PSII core [20].

Among the five extrinsic proteins located at the lumenal side of the diatom PSII core, three (PsbO, PsbU, and PsbV) are commonly found in the cyanobacterial and red algal PSII [1,20,28], and one (PsbQ′) is found in the red algal PSII [20] but not in the cyanobacterial and higher plant PSII. The fifth one, Psb31, is unique to the diatom PSII [29,30]. No homologous sequences of Psb31 are found in other lineages of photosynthetic organisms, but its overall structure is similar to that of PsbQ′ with a characteristic four-helix bundle (Fig. 1C). This may suggest a similar structural and/or functional role of Psb31 with PsbQ′ in protecting the Mn_4CaO_5 cluster from attack by outside reagents [31,32]. The binding site of Psb31 in diatom PSII is similar to the site of PsbTn in PSII–LHCII [2]. Some of the residues at the long C-terminal loop of Psb31 participate in a hydrogen-bond network starting from the Mn_4CaO_5 cluster and mediated by Yz/D1-H190 [1,3,25]. Thus, Psb31 may function to optimize the proton excretion process for water splitting.

The PSII–FCPII dimer supercomplex contains 230 Chls *a*, 58 Chls *c*, 4 pheophytins, 124 Fxs, 20 Bcrs, and 2 Ddxs, as well as some other cofactors involved in the electron transfer chain [25,26]. In addition, 20 digalactosyl diacylglycerol (DGDG), 42 monogalactosyl diacylglycerols (MGDG), 16 sulfoquinovosyl diacylglycerol (SQDG), and 30 phosphatidylglycerol (PG) molecules are found in a PSII–FCPII dimer; these lipids are mostly distributed in the interfaces between subunits (Fig. 1D), suggesting their roles in mediating subunit interactions. These cofactors, together with 70 protein subunits, give rise to a total molecular weight of 1.4 MDa for the PSII–FCPII dimer.

The detailed binding environments of Chls *c* and Fxs in FCP are visualized from a higher resolution (1.8 Å) structure of an isolated FCP subunit from a pennate diatom *P. tricornutum* analyzed by X-ray crystallography [19]. This structure showed that the isolated FCP exists in a dimeric form (Fig. 1E), which is encoded by the *lhcf4* gene. Each FCP monomer contains seven Chls *a*, two Chls *c*, seven Fxs, and one Ddx. Compared with the pigments found in LHCI and LHCII, the total number of Chls in FCP is much smaller (the typical number of Chls in LHCI and LHCII is 13–15), which may be a result of a smaller polypeptide of FCP (18–23 kDa) in comparison with that of LHCII (~ 27 kDa). However, the number of Fxs is much larger (the typical number of Bcr is 3–5 in LHCI and LHCII). This gives rise to a much higher Fx/Chl ratio, which, together with the binding of Chl *c*, provides the basis for the higher ability of FCP to absorb blue-green light as well as to quench the excess energy under a highly fluctuating light environment under water.

Structures of FCP subunits and their interactions

The overall structure of FCP is similar to that of LHCI and LHCII in that they possess three typical membrane-spanning helices (A, B, and C, respectively) (Fig. 2). However, apparent differences are found in some loop regions and the binding sites for pigments between diatom FCPs and LHCI/LHCII subunits of the green lineage organisms (Figs 2 and 3). There are also some differences in the structures of the different FCPs subunits. The main subunit in the FCP-A

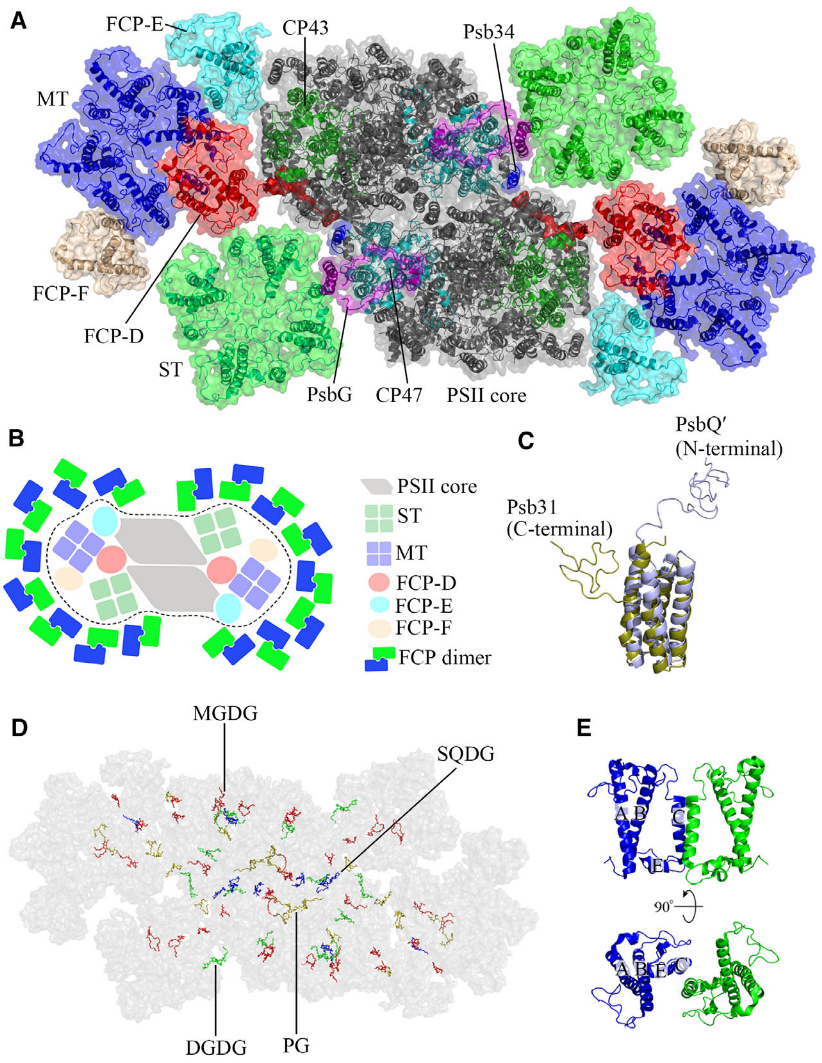

Fig. 1. Overall structures of the PSII–FCPII supercomplex and a FCP dimer. (A) Structure of the PSII–FCPII supercomplex of *C. gracilis* viewed from the stromal side. (B) Schematic representation of the FCP subunit organization within the PSII–FCPII supercomplex shown in panel (A). (C) Comparison of the structures of PsbQ′ (light blue) and Psb31 (yellow orange) from the PSII–FCPII supercomplex of *C. gracilis*. (D) Distributions of lipids in the PSII–FCPII. MGDG, DGDG, SQDG, and PG are colored as red, green, blue, and yellow-orange, respectively. (E) Structure of the FCP dimer from *P. tricornutum*. The figure is generated with PyMOL (https://pymol.org/2/).

tetramer from centric *C. gracilis* has a slightly extended C-terminal region with an additional helix D, resulting in its higher molecular weight of around 22 kDa (Fig. 2A). Four FCP-A proteins are assembled into a ‘head-to-tail’ manner similar to that found in the LHCII trimer, but different from that of Lhcf4 dimer (Fig. 2E-G). The dominant interactions stabilizing the FCP-A tetramer are the tight A-C loop interactions at the stromal side, leading to larger distances between helix D and B-C loop of the adjacent FCP-A monomer at the lumenal side. This results in a bowl shape with its bottom oriented toward the stromal side. The cavity formed at the lumenal side allows coordination of one end group of Fx302/Fx307 and a lipid molecule close to helix D. However, almost no density was clearly observed in the central region of the cavity [25], probably due to the insufficient local resolution. In LHCII, an N-terminal ‘WYGPDR’ motif and a C-terminal Trp residue are found to be crucial for the formation of the trimer by forming hydrogen bonds between adjacent Lhcb monomers [6,33]. However, this motif is absent in FCPs, and characteristic monomer–monomer interactions within the FCP-A tetramer are found between pigment–pigment and pigment–proteins at both stromal and lumenal sides [25,26]. These interactions are different from those found in the dimeric FCP and trimeric LHCII, thus illustrating the different oligomerization states and high flexibility of the protein subunits in the LHC superfamily [6,19,25,26,33], which may be a result of adaptation to fulfill their roles for light harvesting and photoprotection in different light environments.

The FCP-D monomer in the PSII–FCPII supercomplex of *C. gracilis* mediates binding of the MT to CP43 and also holds week interactions with the ST from the adjacent PSII–FCPII monomer unit (Fig. 1A, B). Therefore, this subunit may mediate interactions and energy transfer from MT to the PSII core but also between the two PSII–FCPII monomers in the PSII–FCPII dimer. Interestingly, this FCP-D is an Lhca-like subunit (Fig. 2B) and has the largest molecular weight of around 25 kDa among FCPs; therefore, the nomenclature of pigments in this subunit is similar to those in Lhca2 (Fig. 3C). The structures of FCP-D and LHCI (Lhca2) are found to be similar in that both have three carotenoid-binding sites and a typical Chl–PG–Chl interaction pattern (Fig. 3C,D). Another remarkable feature of FCP-D is its extremely long C-terminal loop (Fig. 2B), which functions as a linker to CP43 [25,26]. Although different from the N-terminal loop of CP29 attached to CP47 at the stromal surface (Fig. 2B,D), the long C-terminal loop of FCP-D enables it to mediate interactions and energy transfer between MT and the PSII core [25,26].

FCP-E and FCP-F monomers are suggested to be Lhcf4-like subunits, among which FCP-E connects the MT and intrinsic PsbZ (Fig. 1A). However, one extra Chl *a*411 is found at the interface between FCP-E and PsbZ that functions to mediate the energy transfer from the MT to PSII core (Fig. 4). FCP-F is located outside of MT and may also interact with ST from the adjacent PSII–FCPII unit, enabling it to mediate energy transfer between the two PSII–FCPII monomers within the PSII–FCPII dimer. These results suggest that the different FCP subunits adopt different sequences and structures to fulfill their different structural and functional roles in PSII–FCPII. No LHCX subunits are identified in the PSII–FCPII supercomplex structure. This may be due to the insufficient resolution and/or loose association of them with PSII, resulting in their loss during purification.

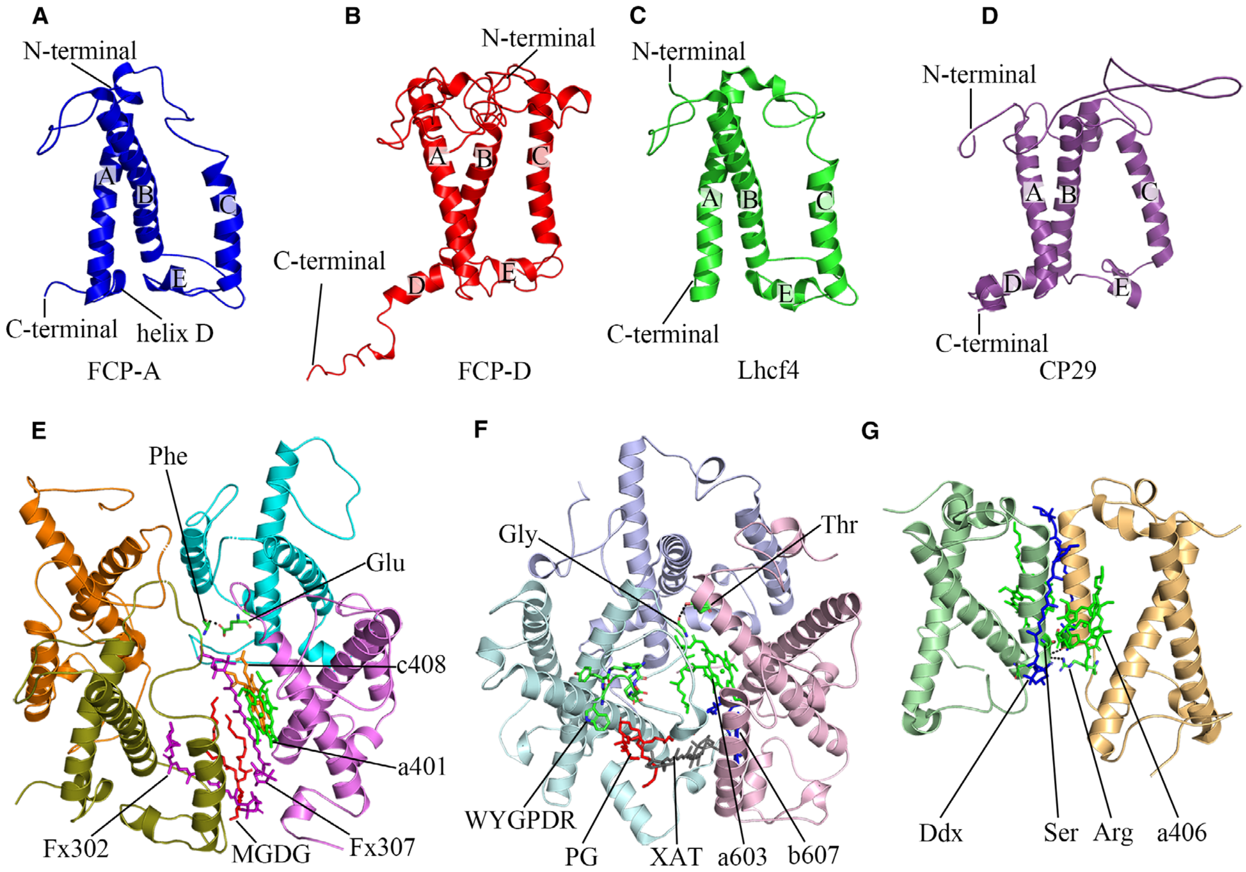

Fig. 2. Structures of the individual subunits of FCP and their oligomerization states in comparison with those of LHC proteins. (A–D) Structures of FCP-A and FCP-D from *C. gracilis* (25), Lhcf4 from *P. tricornutum* (19), and CP29 from *Pisum sativum* (2), respectively. Helices A/B/C/D/E and C/N terminus in each subunit are labeled. (E) Structure of an FCP-A tetramer (25), showing monomer–monomer interactions within the tetramer. Selected pigments and amino acids at a monomer–monomer interface are labeled. (F) Structure of an LHCII trimer (6). Selected pigments and amino acids (including the 'WYGPDR' motif) at a monomer–monomer interface are labeled. (G) Structure of the Lhc4 dimer (19). The figure is generated with PyMOL (https://pymol.org/2/).

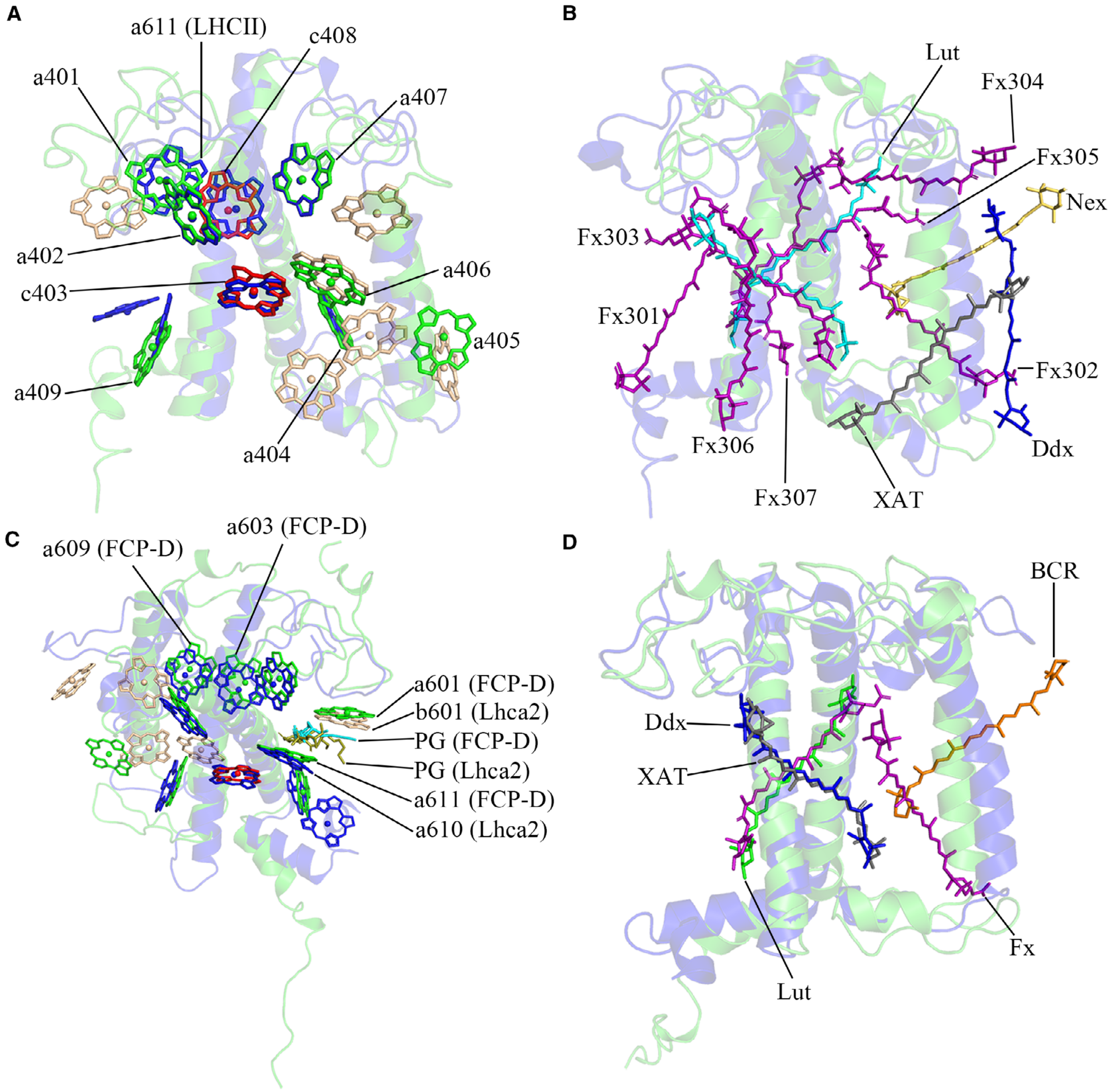

Fig. 3. Comparisons of pigments between FCPs and LHCs. (A) Superposition of the structures of Lhcf4 (green cartoon) and LHCII (marine cartoon). Color for the pigments: Chl *a* in Lhcf4: green; Chl *a* in LHCII: blue; Chl *b* in LHCII: wheat; and Chl *c* in Lhcf4: red. (B) Comparison of carotenoids in Lhcf4 (green cartoon) and LHCII (marine cartoon). Fx, Ddx, lutein (Lut), neoxanthin (Nex), and xanthophyll cycle pigment (XAT) are colored as purple, blue, cyan, yellow, and gray, respectively. (C) Comparisons of chlorophylls bound to FCP-D (green cartoon) and Lhca2 (marine cartoon) [34]. Colors used are as follows: Chl *a* in FCP-D: green; Chl *a* in Lhca2: blue; Chl *b* in Lhca2: wheat; Chl *c* in Lhcf4: red; PG in FCP-D: cyan; and PG in Lhca2: deep olive. (D) Comparisons of carotenoids bound to FCP-D (green cartoon) and Lhca2 (marine cartoon). Fx, Ddx, lutein, Bcr, and XAT are colored as purple, blue, green, orange, and gray, respectively. The figure is generated with PyMOL (https://pymol.org/2/).

Lhcf4 that forms a dimer in the pennate diatom *P. tricornutum* has a shorter helix C without the hydrophilic helix D around its C-terminal loop (Fig. 2C), giving rise to a lower molecular weight of around 18.5 kDa. The large differences in the C-helix between Lhcf4 and those of Lhca and Lhcb enable Lhcf4 to form a homodimer through interactions between the C-helices of two monomers, which are mediated by two strong hydrogen bonds (Arg–Ser) at the stromal side and extensive hydrophobic

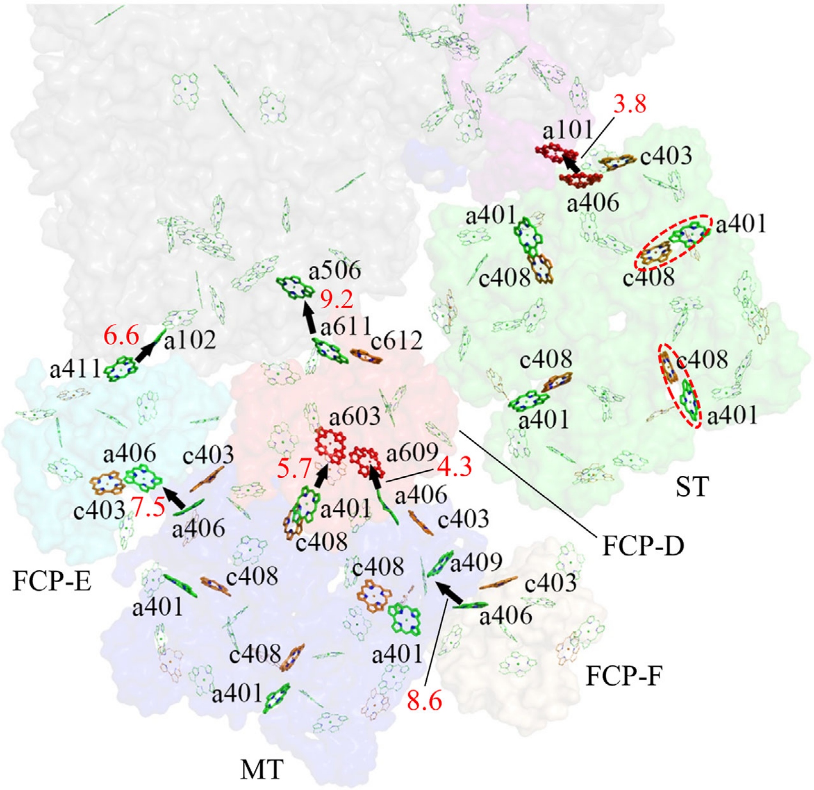

Fig. 4. Chlorophylls bound in one half of the PSII–FCPII supercomplex. Chl *a* and Chl *c* are colored in green and orange, respectively, with a coupled Chl *a* dimer in FCP-D shown in red. Two coupled Chl clusters are encircled by red dashed lines, and arrows show the energy transfer pathways between FCPs and from FCPs to the PSII core. The figure is generated with PyMOL (https://pymol.org/2/).

interactions in the transmembrane region including interactions between two Chls a406, two Ddx308, and a number of hydrophobic residues (Fig. 2G) [19]. The existence and location of the dimer in the centric diatoms are not clear; it may loosely associate at the outside of the PSII–FCPII supercomplex (Fig. 1B), as it is easily dissociated and can be purified separately.

Binding features of chlorophylls and carotenoids in FCPs

The properties of Chl *c* and Fxs in harvesting blue-green light are not only determined by their characteristics of the molecular conjugate systems, but also by the protein environment they bind. High-resolution structure of the Lhcf4 dimer showed that while the binding sites of five Chls *a* in Lhcf4 are similar to those in LHCII, two Chls *a* (a405 and a409) have different orientations and shifted positions as those in LHCII (Fig. 3A) [6,19]. The two Chls *c* bind to sites that are occupied by two Chls *a* in LHCII, whereas the additional Chls *a* and *b* found in LHCII/LHCI are lost in FCP (Fig. 3A). Among the seven Fxs found in Lhcf4, two (Fx303 and Fx305) bind to the same sites as two Luts in LHCII, whereas the remaining five Fxs have different or new binding sites (Fig. 3B). The binding environments for the two end groups of the seven Fxs have different hydrophobicity, which can be categorized into strong, medium, and low [19]. Since the absorption properties of Fx are strongly affected by the polarity of its solution, for example, the higher polarity of the solution, the more redshift of its absorption, the different binding sites confer seven Fxs the ability of absorbing light in different regions, hence enabling FCP to utilize the blue-green light in a wider range [19]. In the crystal structure of the isolated FCP, Fx303 and Fx305 bind to a highly hydrophobic environment, suggesting that these two Fxs will absorb rather 'blue' light. Fx301 and Fx302 have rather polar pockets for their C-end groups and therefore may absorb more 'green' light. Fx306 and Fx307 are close to two Chl *c* molecules, with their C-end rings hydrogen-bonded to the two Chls *c*, suggesting their rather hydrophilic environment and hence red-shifted absorption. Finally, Fx304 lies at the surface of the stromal side with a higher hydrophilic environment for its two end groups; hence, this Fx may have the most red-shifted absorption among the seven Fxs [19].

The most important feature of FCP is its higher ratio of Fxs/Chls, which brings close interactions between Fxs and Chls within the protein matrix. In fact, the higher resolution structure of Lhcf4 revealed that every Chl *a* or Chl *c* is in close interactions with one or more Fxs [19]. This allows efficient energy transfer from Fxs to Chls and vice versa. The reverse energy transfer from Chls to Fxs provides an efficient way of dissipating excess energy, conferring FCP a higher ability of NPQ.

The number of pigments in FCP-A, FCP-E, and FCP-F in the PSII–FCPII supercomplex from the centric *C. gracilis* is 6–9 Chls *a*, 1–3 Chls *c*, and six Fxs [25,26]. Furthermore, FCP-D binds nine Chls *a*, one Chl *c*, two Fxs, and one Ddx. These deviations may result from the different lengths of the FCP polypeptides and their distinctly different locations, which may confer them different roles in energy harvesting and transfer. A similar situation can be found in LHCII and LHCI in the green lineage organisms, where different numbers of Chls *a* and *b*, and slightly different bindings sites of Chls, are also found [2,6,9,21–24].

Energy transfer and dissipation in the PSII–FCPII supercomplex

The Chls in all FCPs are distributed into two layers similar to those found in LHCI and LHCII. In the stromal layer of an FCP-A monomer, six Chls form two Chl *a*-*c*-*a* clusters (Fig. 4), which resemble the Chl 610-611-612 clusters in LHCII with Chl *a*-*a* couplings [6,19,25,26]. Although the lower energy Chl *a* traps cannot be distinguished unambiguously, the marginal Chl *a*401 coupled with Chl *c*408 is suggested as an energy exit (Figs 3A and 4) [19], and Chl a401/a406 may mediate monomer–monomer energy transfer within the tetramer, based on their peripheral locations.

In the PSII–FCPII supercomplex, some intersubunit coupling of Chls between different subunits of FCPs and between FCP and PSII core subunits is found, such as Chl*a*603–*a*609 dimer in FCP-D that is coupled with Chl *a*406 of MT at their interface (Fig. 4), and Chl *a*406 of ST coupled with Chl *a*101 of PsbG at their interface. The edge-to-edge distances of these coupled Chls *a* are < 4.5 Å, suggesting that they may form the red Chl pairs and hence provide important energy transfer pathways from both MT and ST to the PSII core.

Owing to the multiple couplings of Chls at the intersubunit interfaces, multiple possible pathways for the energy transfer among FCP subunits as well as from FCPs to the PSII core can be found at both the stromal and lumenal side layers (Fig. 4) [25,26]. At the stromal side layer, FCP-F transfers its energy to MT, which in turn transfers the energy to CP43 in the PSII core indirectly through FCP-E and FCP-D. FCP-D may also transfer its energy to ST in the adjacent PSII–FCPII unit, providing an energy transfer pathway between two PSII–FCPII monomers within the dimer. At the lumenal side, ST transfers its energy directly to CP47 of the core and also has weak interactions with FCP-D and PSII core of the adjacent PSII–FCPII unit and hence may also transfer its energy to the adjacent PSII–FCPII unit through these interactions [25,26].

In addition to the vast number of Chls, a huge number of Fxs are distributed over the entire FCP units which have close interactions with Chls as mentioned above. These Fxs help diatoms to absorb blue-green light and transfer them to the Chls efficiently under low-light conditions, and to dissipate excess energy by accepting energy from nearby Chls under strong light conditions. The combination of this complicated network of Chls and Fxs greatly contributes to the successful survival of diatoms in the aquatic environment which has a highly fluctuating light environment due to the continuous flow and mixing of the water layers. The possible energy transfer and dissipating pathways discussed here, however, are based primarily on the pigment–pigment distances. More detailed analysis of the pathways and mechanisms will require a more extensive treatment employing calculations by taking into considerations of both distances and dipole–dipole interactions of the pigments, which are now possible based on the structures solved.

Acknowledgements

This work was supported by the National Key R&D Program of China (2017YFA0503700 and 2019YFA0906300), a CAS Key Research Program for Frontier Science (QYZDY-SSW-SMC003), and the National Natural Science Foundation of China (31600191).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

WW, SZ, and J-RS wrote the manuscript, and XP, TK, and SFS contributed to the discussion and revision of the manuscript.

References

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

2 Su X, Ma J, Wei X, Cao P, Zhu D, Chang W, Liu Z, Zhang X & Li M (2017) Structure and assembly

mechanism of plant C2S2M2-type PSII-LHCII supercomplex. *Science* **357**, 815–820.

3 Shen JR (2015) Structure of photosystem II and the mechanism of water oxidation in photosynthesis. *Annu Rev Plant Biol* **66**, 23–48.

4 Wei X, Su X, Cao P, Liu X, Chang W, Li M, Zhang X & Liu Z (2016) Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. *Nature* **534**, 69–74.

5 Shen L, Huang Z, Chang S, Wang W, Wang J, Kuang T, Han G, Shen JR & Zhang X (2019) Structure of a C2S2M2N2-type PSII–LHCII supercomplex from the green alga *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci USA* **116**, 21246–21255.

6 Liu Z, Yan H, Wang K, Kuang T, Zhang J, Gui L, An X & Chang W (2004) Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428**, 287–293.

7 Goss R & Lepetit B (2015) Biodiversity of NPQ. *J Plant Physiol* **172**, 13–32.

8 Croce R & Amerongen HV (2014) Natural strategies for photosynthetic light harvesting. *Nat Chem Biol* **10**, 492.

9 Cao P, Su X, Pan X, Liu Z, Chang W & Li M (2018) Structure, assembly and energy transfer of plant photosystem II supercomplex. *Biochim Biophys Acta* **1859**, 633–644.

10 Falkowski PG & Chen YB (2003) Photoacclimation of light harvesting systems in eukaryotic algae. In Light-harvesting antennas in photosynthesis (Green BR & Parson WW, eds), pp. 423–447. Kluwer Academic Publishers, London.

11 Yu LJ, Suga M, Wang-Otomo ZY & Shen JR (2018) Structure of photosynthetic LH1-RC super-complex at 1.9 Å resolution. *Nature* **556**, 209–213.

12 Yu LJ, Suga M, Wang-Otomo ZY & Shen JR (2018) Novel features of LH1-RC from *Thermochromatium tepidum* revealed from its atomic resolution structure. *FEBS J* **285**, 4359–4366.

13 Gisriel C, Sarrou I, Ferlez B, Golbeck JH, Redding KE & Fromme R (2017) Structure of a symmetric photosynthetic reaction center-photosystem. *Science* **357**, 1021–1025.

14 Field CB, Behrenfeld MJ, Randerson JT & Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237–240.

15 Gundermann K & Büchel C (2014) Structure and functional heterogeneity of fucoxanthin-chlorophyll proteins in diatoms. In The Structural Basis of Biological Energy Generation (Hohmann-Marriott MF, ed), pp. 21–37. Springer, Dordrecht, Heidelberg, New York, London.

16 Büchel C (2014) Fucoxanthin-chlorophyll-proteins and non-photochemical fluorescence quenching of diatoms. In Non-Photochemical Quenching and Energy Dissipation in Plants (Demmig-Adams B, Garab G, Adams W & Govindjee G, eds), pp. 259–275. Springer, Dordrecht, Heidelberg, New York, London.

17 Zapata M, Garrido JL & Jeffrey SW. Chlorophyll c pigments: current status. In Chlorophylls and Bacteriochlorophylls, Advances in Photosynthesis and Respiration (Grimm B, Porra RJ, Rüdiger W & Scheer H eds), pp. 38–53. Springer, Dordrecht, Heidelberg, New York, London.

18 Roding A, Boekema E & Büchel C (2018) The structure of FCPb, a light-harvesting complex in the diatom *Cyclotella meneghiniana*. *Photosynth Res* **135**, 203–211.

19 Wang W, Yu LJ, Xu C, Tomizaki T, Zhao S, Umena Y, Chen X, Qin X, Xin Y, Suga M *et al.* (2019) Structural basis for blue-green light harvesting and energy dissipation in diatoms. *Science* **363**, eaav0365.

20 Ago H, Adachi H, Umena Y, Tashiro T, Kawakami K, Kamiya N, Tian L, Han G, Kuang T, Liu Z *et al.* (2016) Novel features of eukaryotic photosystem II revealed by its crystal structure analysis from a red alga. *J Biol Chem* **291**, 5676–5687.

21 Pi X, Tian L, Dai HE, Qin X, Cheng L, Kuang T, Sui F & Shen JR (2018) Unique organization of photosystem I-light harvesting supercomplex revealed by cryo-EM from a red alga. *Proc Natl Acad Sci USA* **115**, 4423–4428.

22 Su X, Ma J, Pan X, Zhao X, Chang W, Liu Z, Zhang X & Li M (2019) Antenna arrangement and energy transfer pathways of a green algal photosystem-I-LHCI supercomplex. *Nat Plants* **5**, 273–281.

23 Qin X, Pi X, Wang W, Han G, Zhu L, Liu M, Cheng L, Shen JR, Kuang T & Sui SF (2019) Structure of a green algal photosystem I in complex with a large number of light-harvesting complex I subunits. *Nat Plants* **5**, 263–272.

24 Suga M, Ozawa SI, Yoshida-Motomura K, Akita F, Miyazaki N & Takahashi Y (2019) Structure of the green algal photosystem I supercomplex with a decameric light-harvesting complex I. *Nat Plants* **5**, 626–636.

25 Pi X, Zhao S, Wang W, Liu D, Xu C, Han G, Kuang T, Sui SF & Shen JR (2019) The pigment-protein network of a diatom photosystem II-light-harvesting antenna supercomple. *Science* **365**, eaax0446.

26 Nagao R, Kato K, Suzuki T, Ifuku K, Uchiyama I, Kashino Y, Dohmae N, Akimoto S, Shen JR, Miyazaki N *et al.* (2019) Structural basis for energy harvesting and dissipation in a diatom PSII-FCPII. *Nat Plants* **5**, 890–901.

27 Nagao R, Tomo T, Narikawa R, Enami I & Ikeuchi M (2013) Light-independent biosynthesis and assembly of the photosystem II complex in the diatom *Chaetoceros gracilis*. *FEBS Lett* **587**, 1340–1345.

28 Suga M, Akita F, Hirata K, Ueno G, Murakami H, Nakajima Y, Shimizu T, Yamashita K, Yamamoto M, Ago H *et al*. (2015) Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99–103.

29 Okumura A, Nagao R, Suzuki T, Yamagoe S, Iwai M, Nakazato K & Enami I (2008) A novel protein in Photosystem II of a diatom *Chaetoceros gracilis* is one of the extrinsic proteins located on lumenal side and directly associates with PSII core components. *Biochim Biophys Acta* **1777**, 1545–1551.

30 Nagao R, Suga M, Niikura A, Okumura A, Koua FH, Suzuki T, Tomo T, Enami I & Shen JR (2013) Crystal structure of Psb31, a novel extrinsic protein of photosystem II from a marine centric diatom and implications for its binding and function. *Biochemistry* **52**, 6646–6652.

31 Shen JR, Ikeuchi M & Inoue Y (1997) Analysis of the psbU gene encoding the 12 kDa extrinsic protein of photosystem II and studies on its role by deletion mutagenesis in *Synechocystis* sp. PCC 6803. *J Biol Chem* **272**, 17821–17826.

32 Shen JR, Qian M, Inoue Y & Burnap RL (1998) Functional characterization of *Synechocystis* sp. PCC 6803 ΔpsbU and ΔpsbV mutants reveals important roles of cytochrome c-550 in cyanobacterial PSII. *Biochemistry* **37**, 1551–1558.

33 Barros T & Kühlbrandt W (2009) Crystallisation, structure and function of plant light-harvesting complex II. *Biochim Biophys Acta* **1787**, 753–772.

34 Qin X, Suga M, Kuang T & Shen JR (2015) Structural basis for the energy transfer pathways in plant PSI-LHCI super-complex. *Science* **348**, 989–995.