Supporting Information for

Structure of a $C_2S_2M_2N_2$-type PSII-LHCII supercomplex from the green alga *Chlamydomonas reinhardtii*

Liangliang Shen, Zihui Huang, Shenghai Chang, Wenda Wang, Jingfen Wang, Tingyun Kuang, Guangye Han, Jian-Ren Shen, Xing Zhang

Corresponding authors

Xing Zhang
xzhang1999@zju.edu.cn

Jian-Ren Shen
shen@cc.okayama-u.ac.jp

Guangye Han
hanguangye@ibcas.ac.cn.

This PDF file includes:
Figs. S1 to S10
Table S1 to S2

www.pnas.org/cgi/doi/10.1073/pnas.1912462116

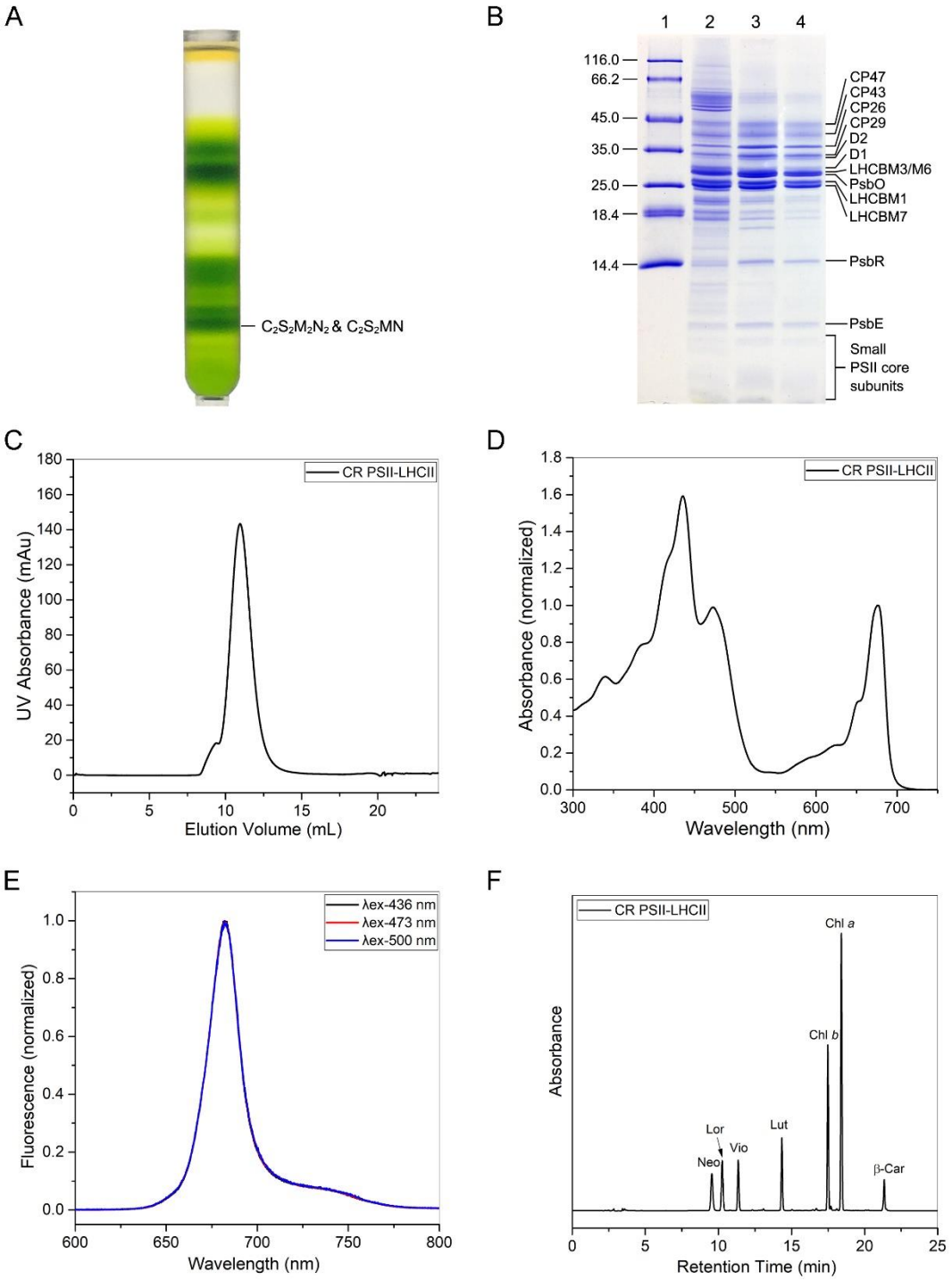


Fig. S1. Purification and characterization of the PSII-LHCII supercomplex from *Chlamydomonas reinhardtii*. (*A*) Isolation of PSII-LHCII by sucrose density gradient centrifugation. The lowest band labeled $C_2S_2M_2N_2$ & C_2S_2MN was collected and used for single particle analysis in this study. (*B*) SDS-PAGE analysis of the purified PSII-LHCII from *C. reinhardtii*. Lane 1: marker; lane 2: thylakoid membranes (5 μg Chl); lane 3: PSII-LHCII after sucrose density gradient centrifugation (5 μg Chl); lane 4: PSII-LHCII after size-exclusion chromatography (5 μg Chl). (*C*) Elution profile of the PSII-

LHCII supercomplex from size-exclusion chromatography. Elution was performed at a flow rate of 100 µl min^{-1} at 4°C and monitored by absorption at 280 nm. (*D*) Room-temperature absorption spectrum of PSII-LHCII obtained from size-exclusion chromatography. (*E*) Room temperature fluorescence emission spectra of PSII-LHCII after size-exclusion chromatography upon preferential excitation of Chl *a* at 436 nm, Chl *b* at 473 nm and carotenoids at 500 nm. The maximum emissions of all spectra were at 682 nm. Overlapping of the emission spectra indicates that there are no free pigments in the purified PSII-LHCII sample. (*F*) Analysis of the pigment composition of PSII-LHCII after size-exclusion chromatography by HPLC. Based on the characteristic absorption spectrum and elute time of each fraction, seven major pigment peaks were identified which are neoxanthin (Neo), loroxanthin (Lor), violaxanthin (Vio), lutein (Lut), chlorophyll *b* (Chl *b*), chlorophyll *a* (Chl *a*) and β-carotene (β-Car), respectively.

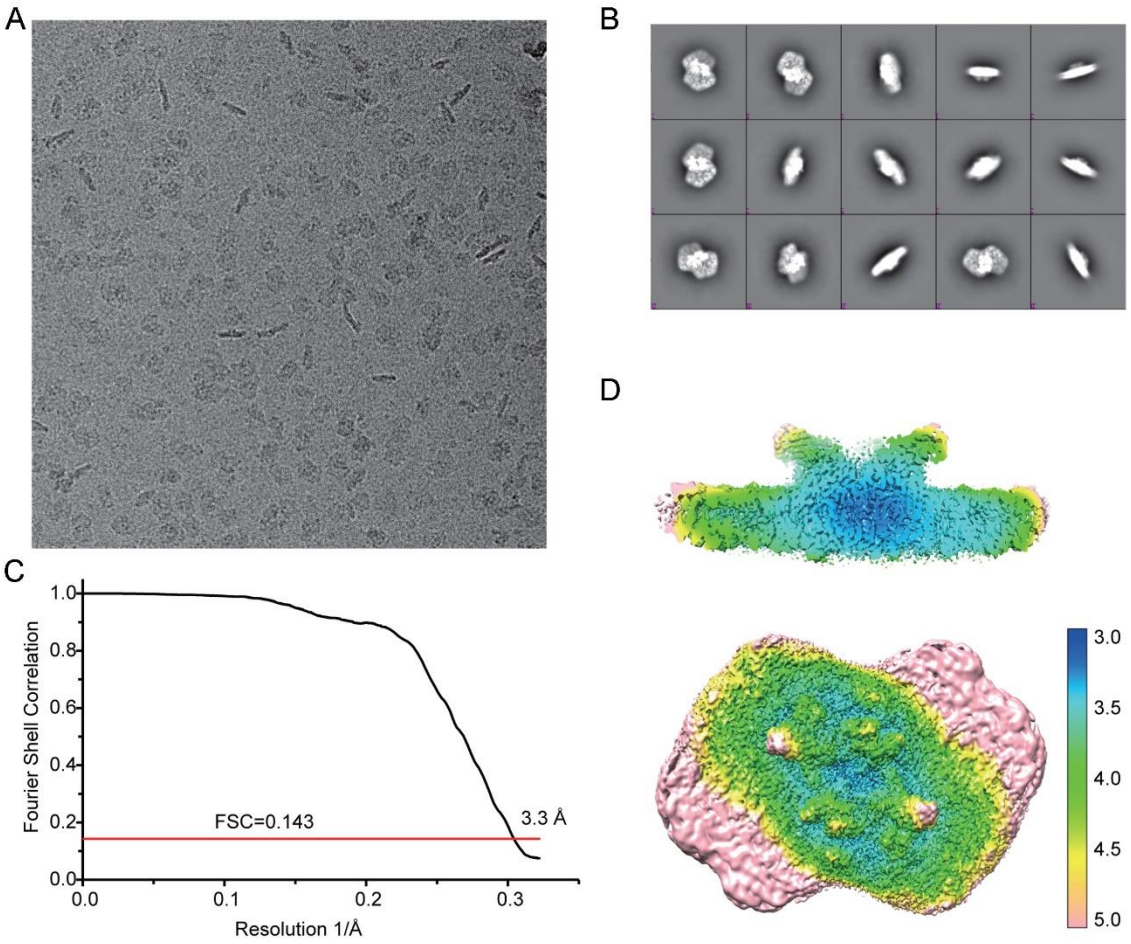


Fig. S2. Cryo-EM images and the cryo-EM map quality. (*A*) A representative of cryo-EM micrographs of the $C_2S_2M_2N_2$-type PSII-LHCII supercomplex from *C. reinhardtii*. (*B*) Representative 2D class average images. (*C*) The gold standard FSC curves of the final 3D reconstruction of the PSII core region. (*D*) Local resolution distributions of the $C_2S_2M_2N_2$-type PSII-LHCII generated with Relion.

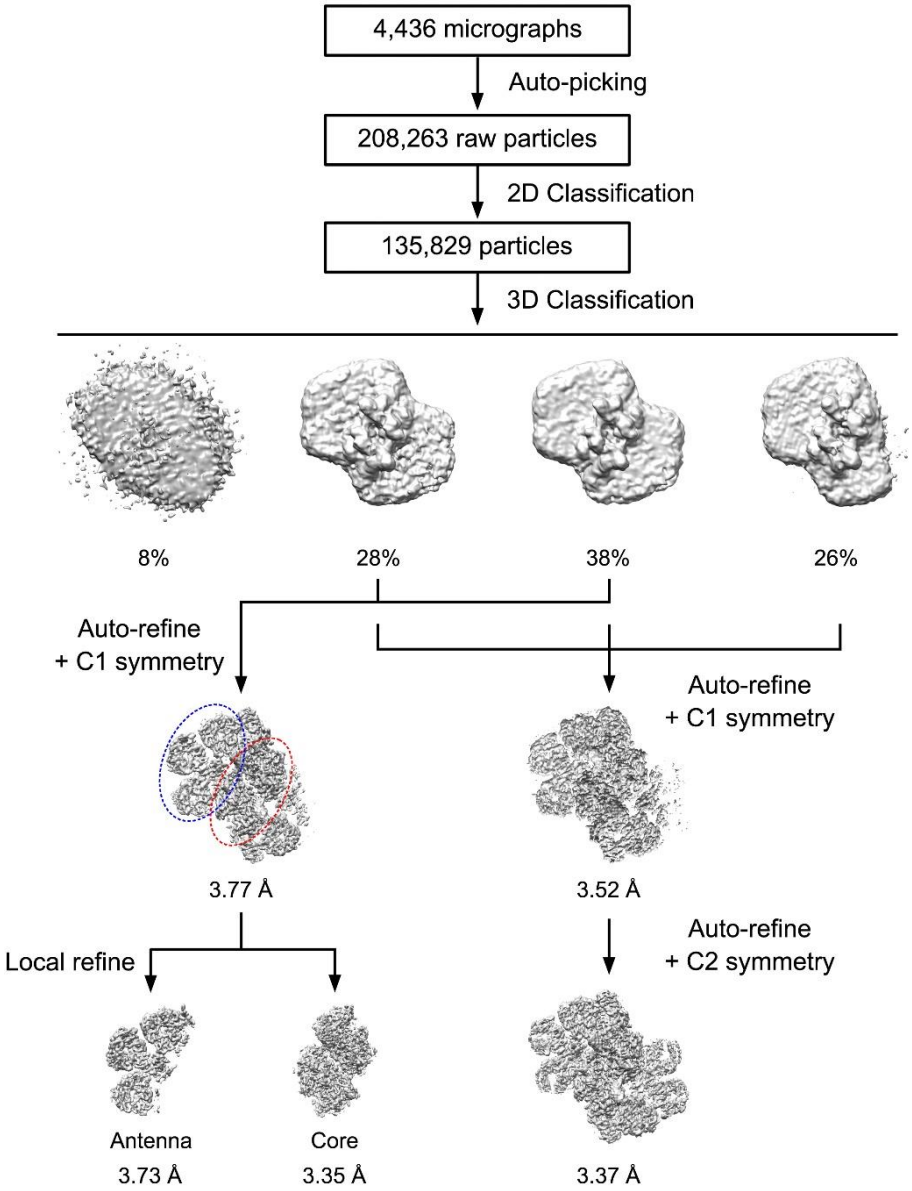


Fig. S3. Flow chart of the cryo-EM data analysis. From 4,436 drift-corrected micrographs, 208,263 raw particles were picked up and used for 2D classifications. From the results of 2D classifications, 50 classes and 135,829 particles were selected for 3D classification which generated 4 classes. The left-side class does not correspond to any of the PSII-LHCII particles, and thus were not used in the subsequent processing. The rest three classes contained ~124,456 particles and were further selected for 3D refinement using C1 and then C2 symmetry which yielded maps with resolution of 3.52 Å and 3.37 Å, respectively. On the other hand, 89,018 particles from two assembled classes (in the middle) were also selected for 3D refinement with a C1 symmetry and yielded a map of 3.77 Å. To improve the resolution, local-refinement of the regions of the core (red dashed-line circle) and the light-harvesting antenna complexes II (LHCII) (blue dashed-line circle) were performed, which yielded maps with resolutions of 3.35 Å and 3.73 Å, respectively.

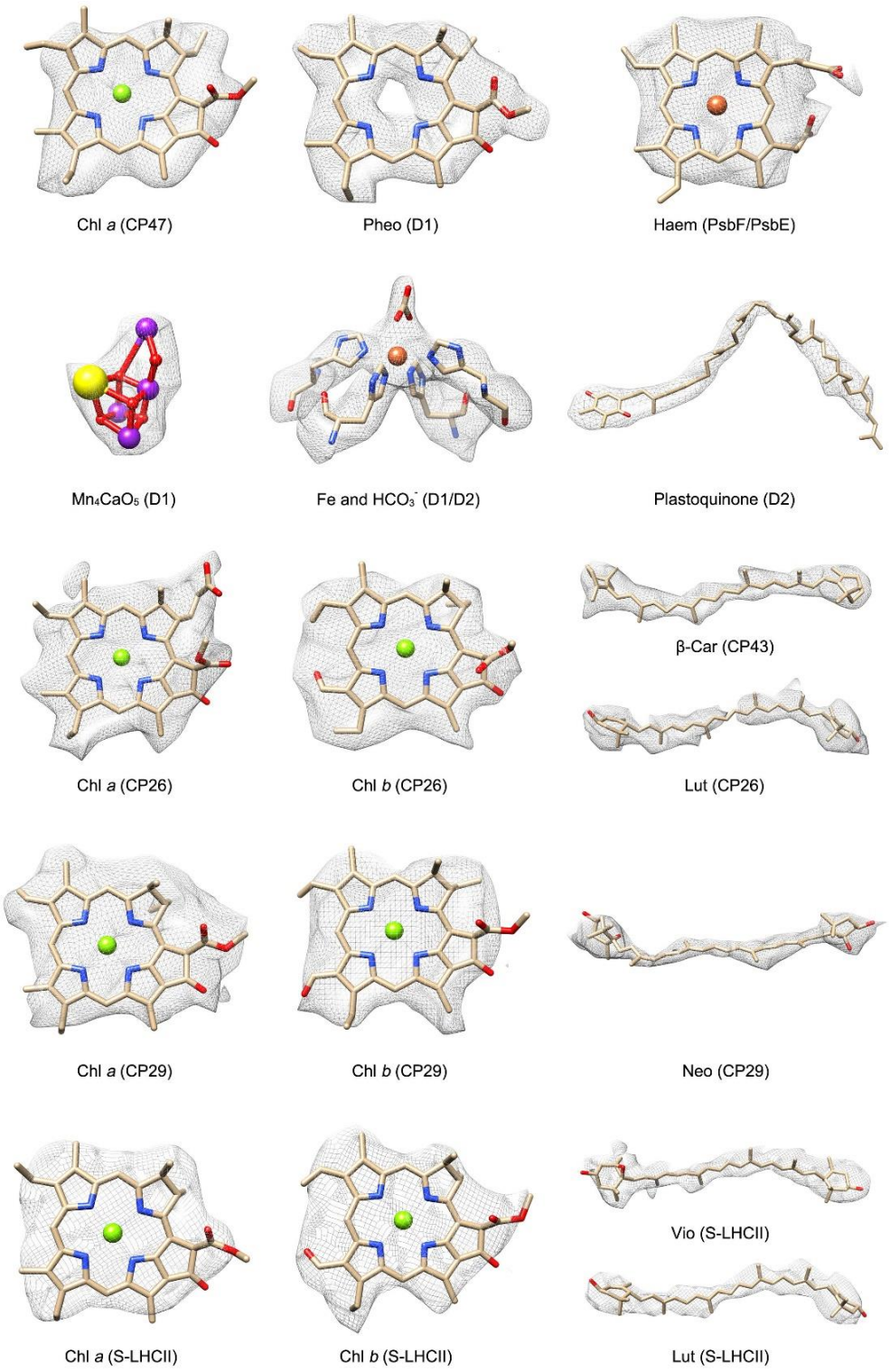


Fig. S4. Cryo-EM densities of various cofactors bound in the $C_2S_2M_2N_2$-type PSII-LHCII supercomplex of *C. reinhardtii*. The density is shown as grey meshes and all cofactors are shown as stick models. Fe and Mg are shown as orange and green spheres. The Ca, Mn, and O of the oxygen-evolving complex are shown as yellow, purple and red, respectively.

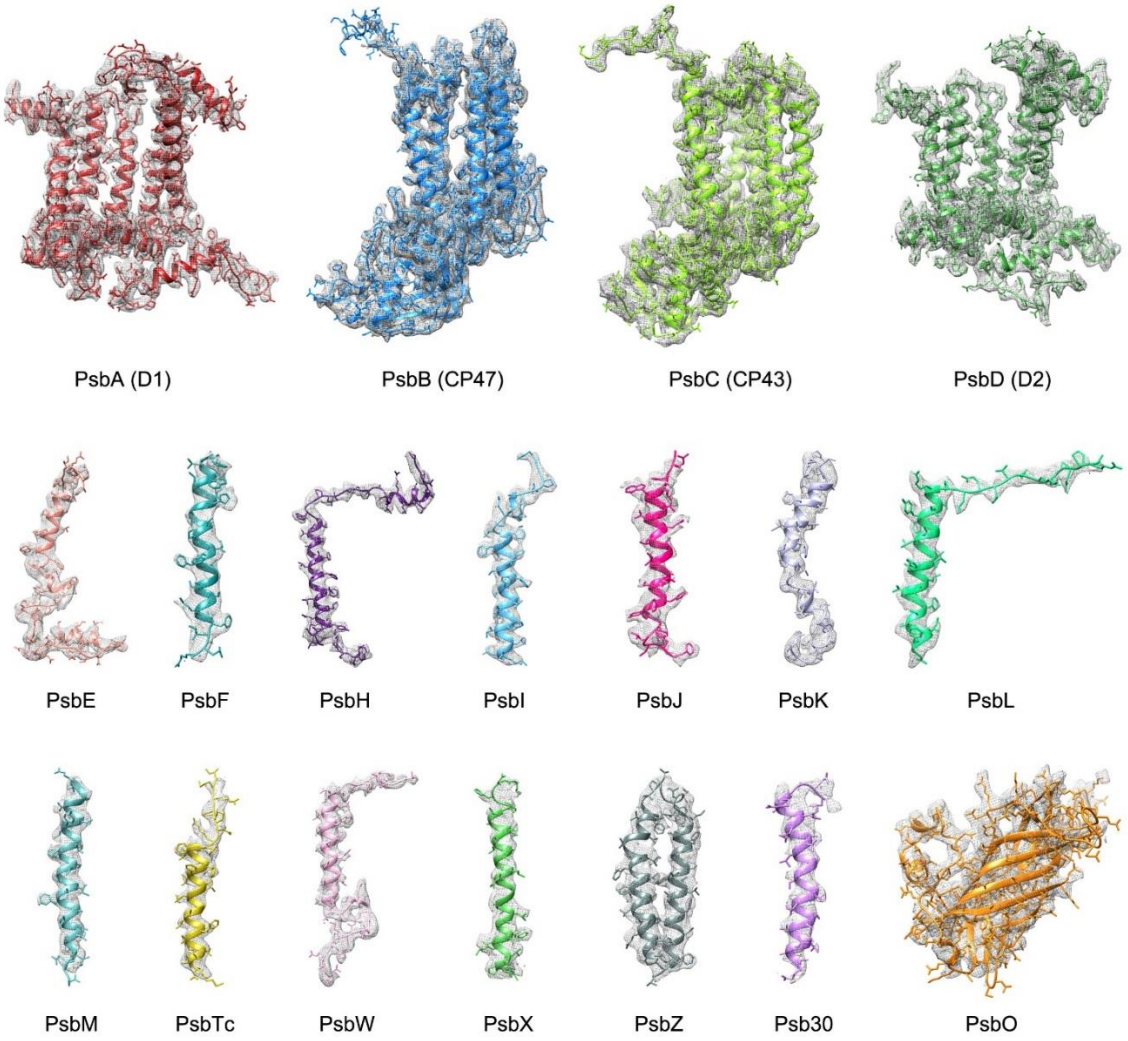


Fig. S5. Cryo-EM densities and structural models of the PSII core intrinsic and extrinsic subunits. The small intrinsic subunits and extrinsic subunit are shown as mixed cartoon/stick and colored the same as in Fig. 1*C*. The cryo-EM density map of each subunit is depicted in gray meshes.

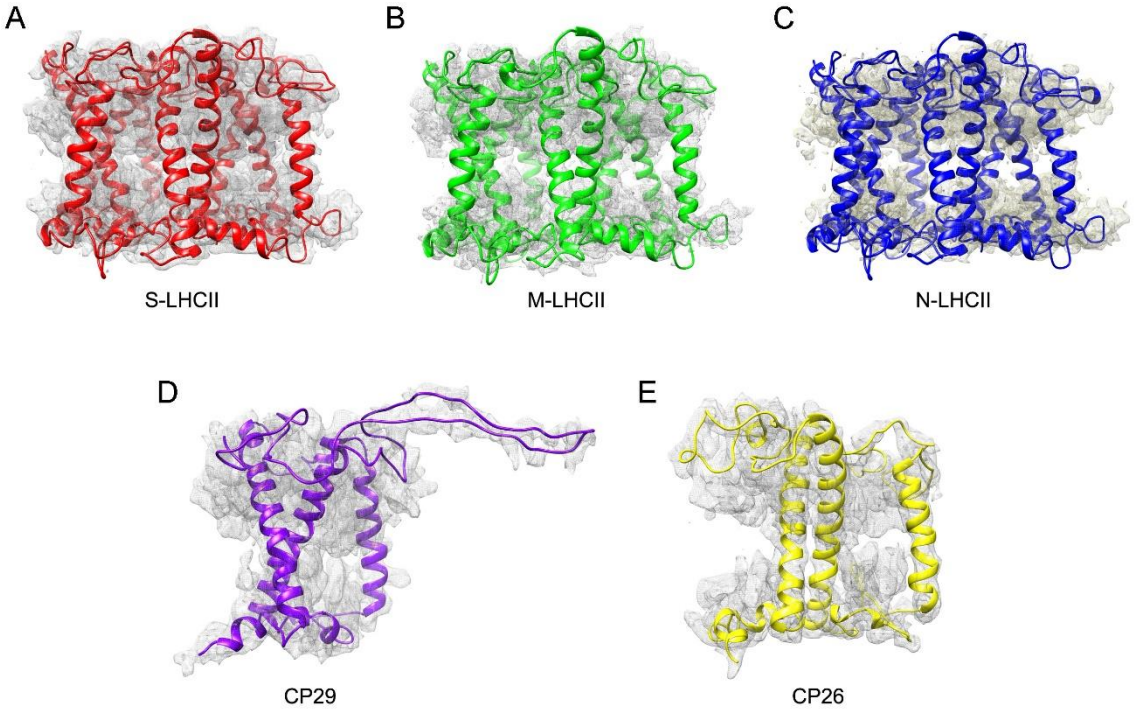


Fig. S6. Cryo-EM density and the structures of LHCII trimers, CP29 and CP26 of *C. reinhardtii*. (*A*) S-LHCII. (*B*) M-LHCII. (*C*) N-LHCII. (*D*) CP29. (*E*) CP26. In each antenna subunit, only the main chains are shown in cartoons with the same color as in Fig. 1*A*.

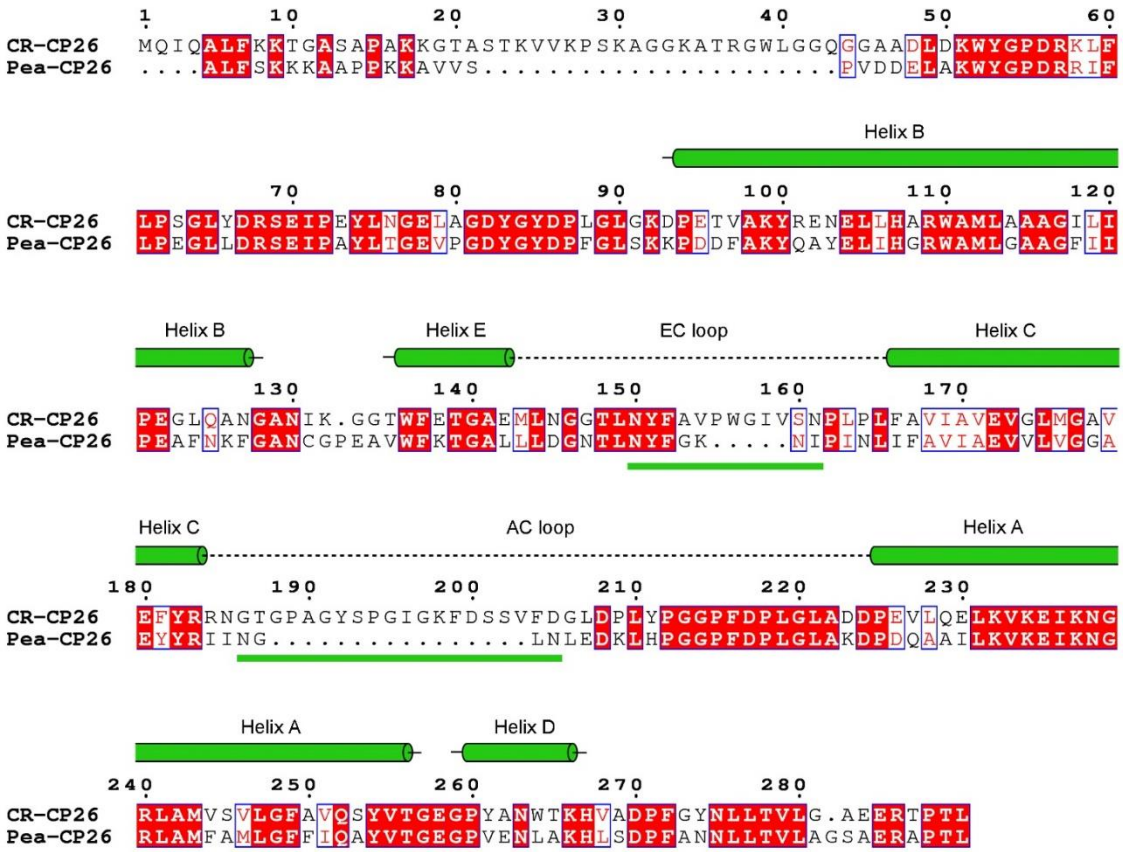


Fig. S7. Sequence alignment of CP26 between pea (pdb code: 5XNL) and *C. reinhardtii*. The information of secondary structures shown above sequences are taken from the cryo-EM structure of CP26 in the $C_2S_2M_2N_2$-type PSII-LHCII supercomplex of *C. reinhardtii*. Helices are represented by green cylinders, and loops are denoted by dashed lines. Sequences corresponding to the two extra loops in CP26 of *C. reinhardtii* are marked underneath with green solid lines.

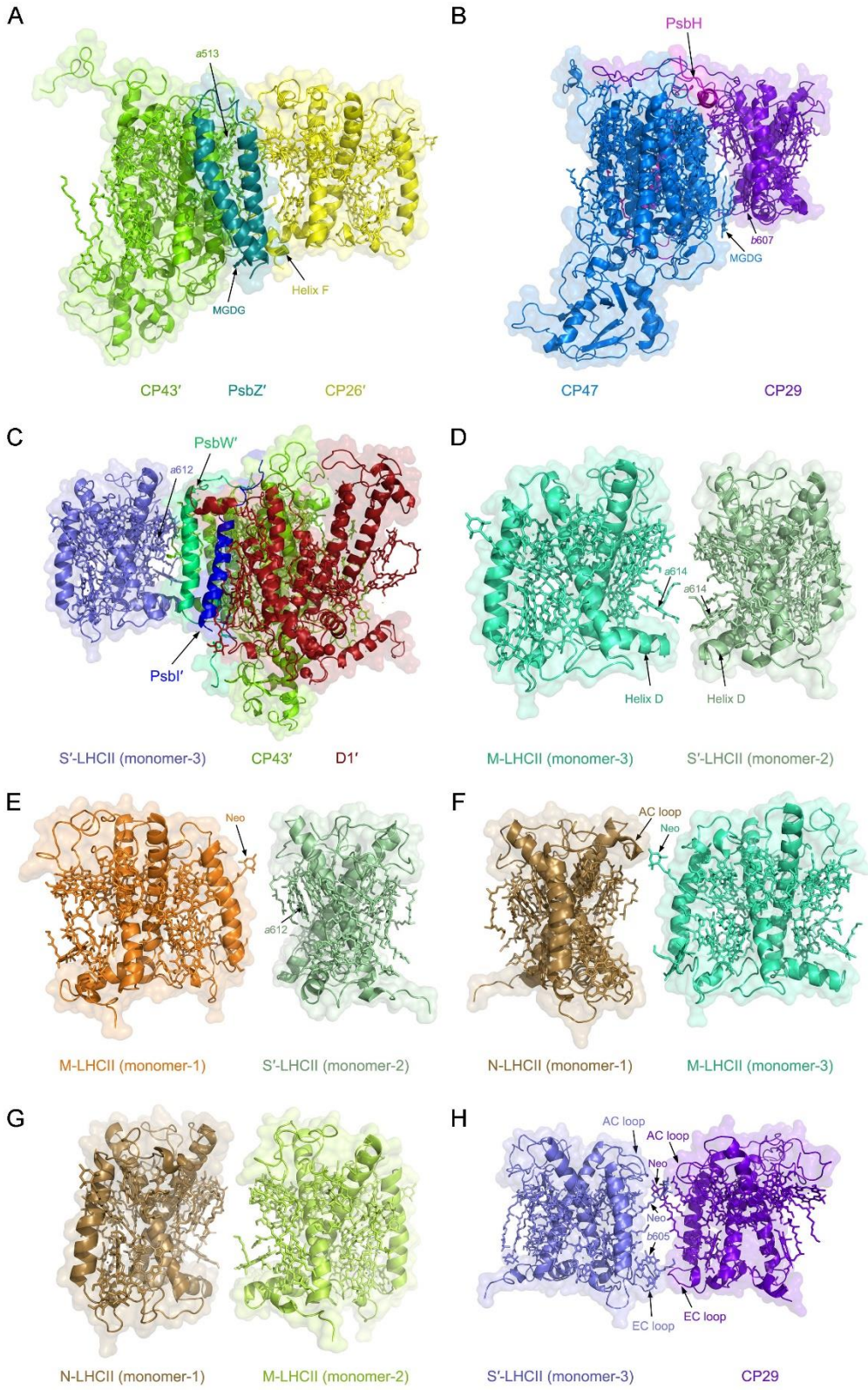


Fig. S8. Antenna-antenna and antenna-core interactions in the $C_2S_2M_2N_2$-type PSII-LHCII supercomplex of *C. reinhardtii* shown in surface and cartoon models. (*A*) Interactions between CP26′

and CP43′ mediated by PsbZ′. (*B*) Interactions between CP47 and CP29. (*C*) Interactions between S′-LHCII (monomer-3) and PSII core subunits mediated by PsbW′ and PsbI′. (*D*) Interactions between M-LHCII (monomer-3) and S′-LHCII (monomer-2). (*E*) Interactions between M-LHCII (monomer-1) and S′-LHCII (monomer-2). (*F*) Interactions between N-LHCII (monomer-1) and M-LHCII (monomer-3). (*G*) Interactions between N-LHCII (monomer-1) and M-LHCII (monomer-2). (*H*) Interactions between S′-LHCII (monomer-3) and CP29. Subunits are colored as that in Fig. 2*A*.

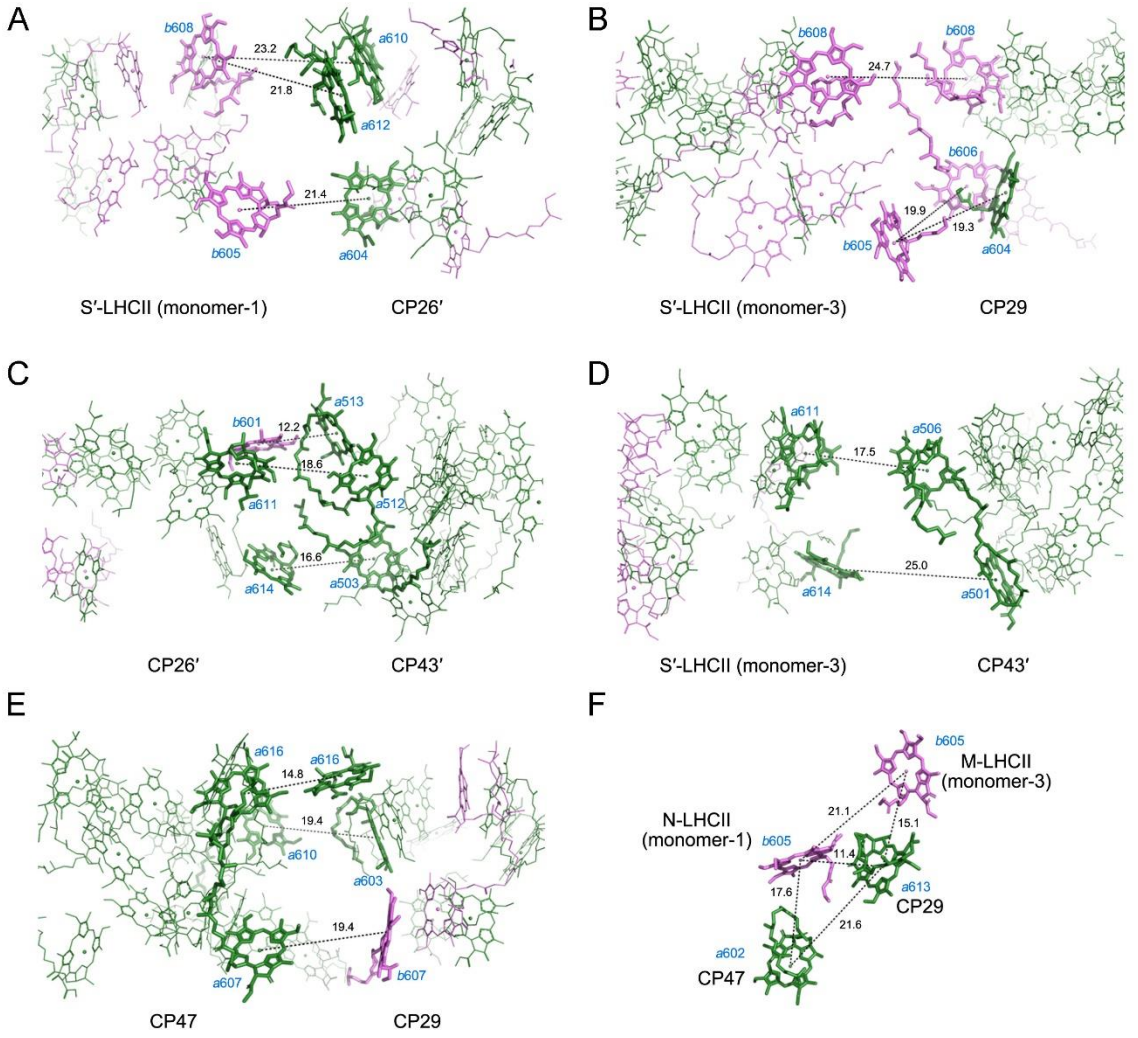
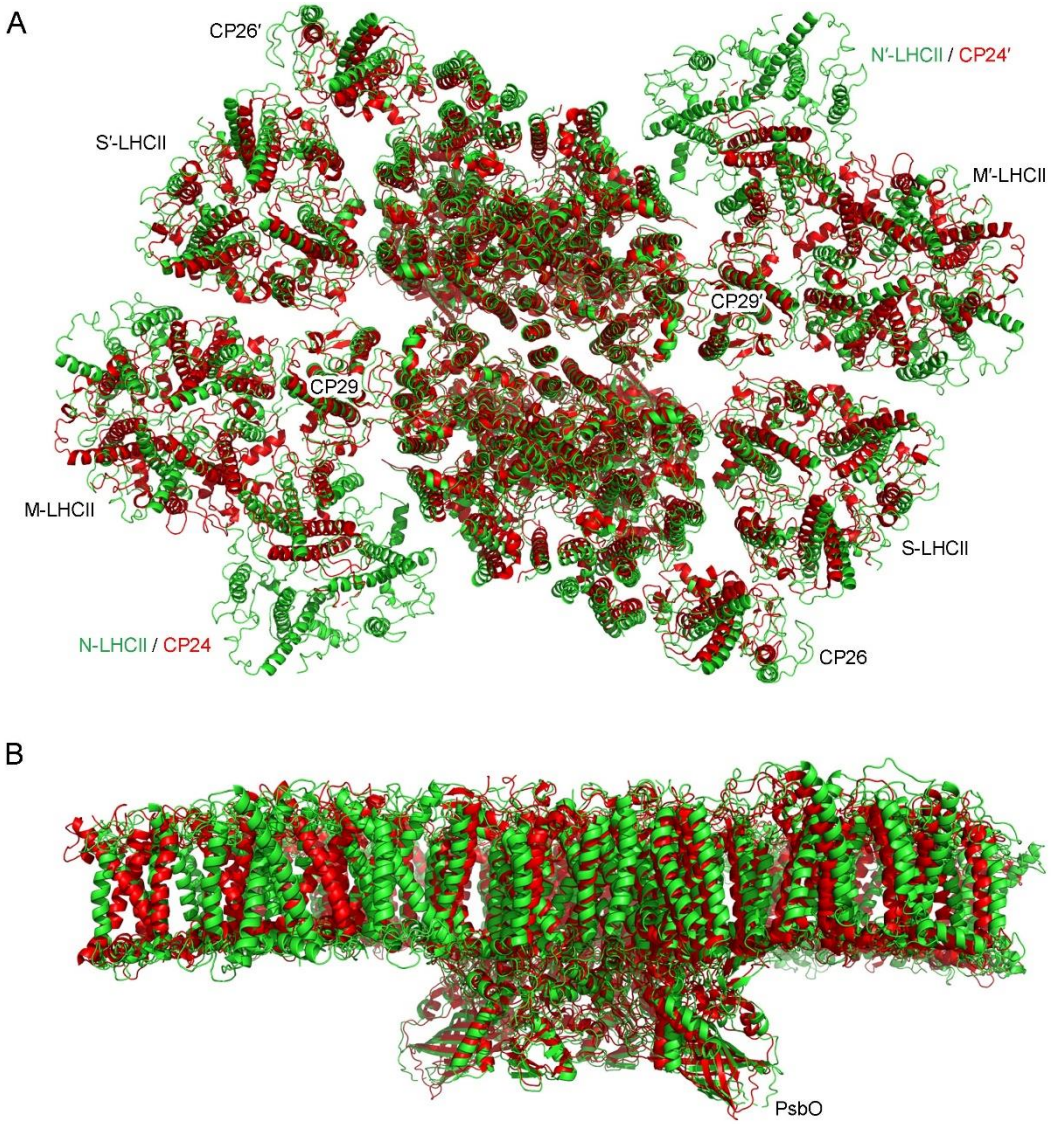


Fig. S9. The interfacial chlorophylls between antenna complexes in the $C_2S_2M_2N_2$-type PSII-LHCII supercomplex of *C. reinhardtii*. (*A*-*E*) The interfacial chlorophylls between S′-LHCII (monomer-1) and CP26′ (*A*), S′-LHCII (monomer-3) and CP29 (*B*), CP26′ and CP43′ (*C*), S′-LHCII (monomer-3) and CP43′ (*D*), CP47 and CP29 (*E*). (*F*) A chlorophyll cluster formed in the interface between M-LHCII (monomer-3), N-LHCII (monomer-1), CP47 and CP29. Chl *a* and Chl *b* are colored in green and violet, respectively. Chlorophylls involved in energy transfer are highlighted as sticks and labeled with blue. The Mg-to-Mg distances (Å) between two adjacent interfacial chlorophylls are indicated in black.


A
CP26′
N′-LHCII / CP24′
S′-LHCII
M′-LHCII
CP29′
CP29
M-LHCII
S-LHCII
N-LHCII / CP24
CP26
B
PsbO

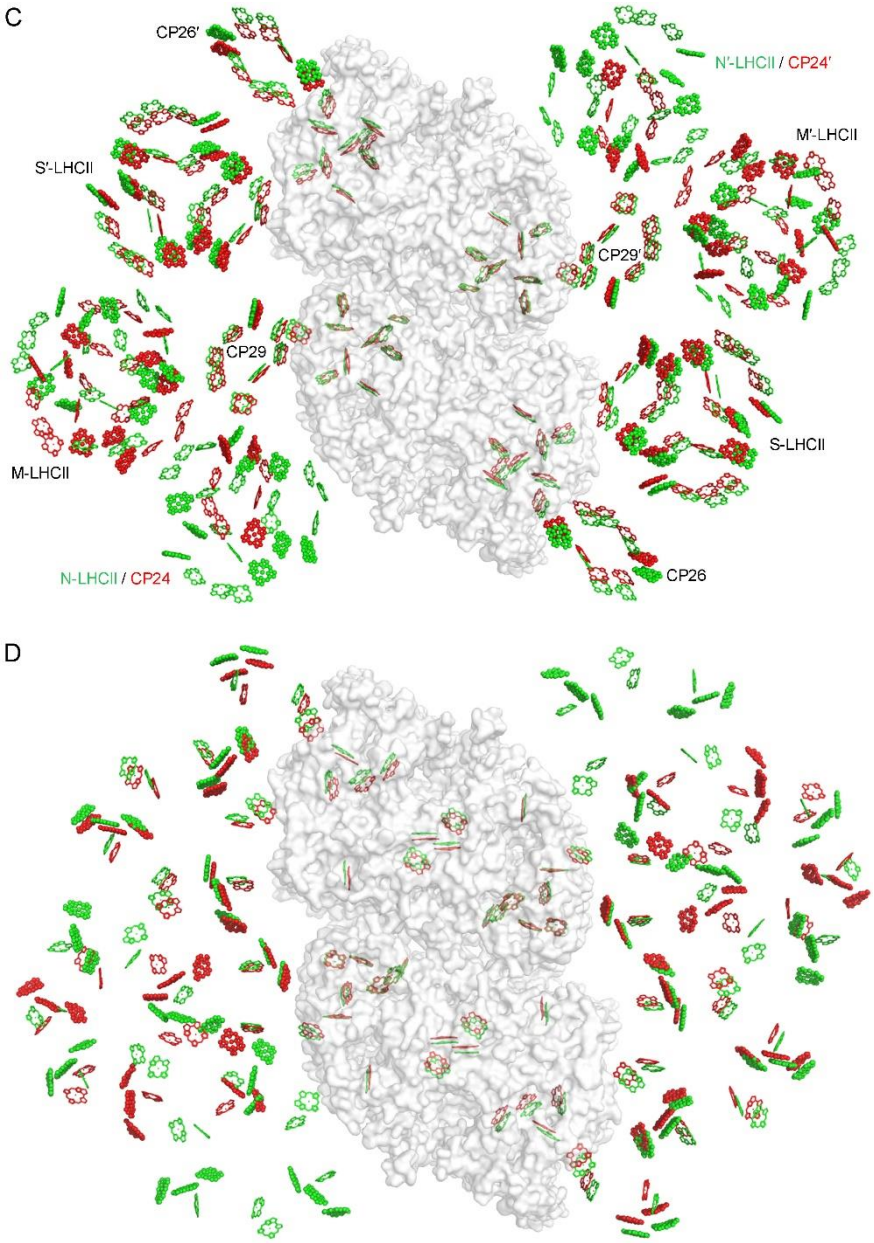


Fig. S10. Structural comparison between the $C_2S_2M_2N_2$-type PSII-LHCII from the green alga (*C. reinhardtii*) and the $C_2S_2M_2$-type PSII-LHCII from a higher plant (pea). (*A* and *B*), Superposition of the protein structure of the $C_2S_2M_2N_2$-type PSII-LHCII from *C. reinhardtii* with $C_2S_2M_2$-type PSII-LHCII from pea with a view from the stromal side (*A*) and along the membrane plane (*B*), respectively. (*C* and *D*), Superposition of Chls of the $C_2S_2M_2N_2$-type PSII-LHCII from *C. reinhardtii* with $C_2S_2M_2$-type PSII-LHCII from pea at the stromal side (*C*) and at the lumenal side (*D*), respectively, with a view from the stromal side. Chls *a* are shown as sticks and Chls *b* are shown as spheres. Color codes: green of $C_2S_2M_2N_2$-type PSII-LHCII (green alga); red of $C_2S_2M_2$-type PSII-LHCII (higher plant). PDB code: 5XNL of pea.

Table S1. Statistics of cryo-EM data and structural analysis of the $C_2S_2M_2N_2$-type PSII-LHCII supercomplex refined at 3.37 Å resolution.

	PSII-LHCII (EMDB-9957) (PDB 6KAF)
Data collection and processing	
Magnification	38,256 ×
Voltage (kV)	300
Electron exposure (e–/Å^2)	50
Defocus range (μm)	-1.5 to -3.0
Pixel size (Å)	1.307
Symmetry imposed	C2
Initial particle images (no.)	208,263
Final particle images (no.)	89,018
Map resolution (Å)	3.37
FSC threshold	0.143
Refinement	
Initial model used (PDB code)	5XNL
Map sharpening B factor (Å^2)	98.0
Model composition	
Non-hydrogen atoms	104,807
Protein residues	74,266
Ligands	30,541
B factors (Å^2)	
Protein	95.36
Ligand	98.96
R.m.s. deviations	
Bond lengths (Å)	0.021
Bond angles (^{o})	1.68
Validation	
MolProbity score	2.31
Clashscore	17.12
Poor rotamers (%)	0.66
Ramachandran plot	
Favored (%)	88.98
Allowed (%)	10.69
Disallowed (%)	0.33

Table S2. Cofactors found in each subunit of the $C_2S_2M_2N_2$-type PSII-LHCII supercomplex.

Subunit	Traced residues	Chlorophylls	Carotenoids	Lipids	Others
PsbA (D1)	333 (11-343)	4 Chl *a* 2 Pheo	1 BCR	2 MGDG 2 SQDG	1 Mn_4CaO_5 1 Fe ion 1 plastoquinone (partially occupied)
PsbB (CP47)	503 (2-504)	16 Chl *a*	3 BCR	2 MGDG 2 SQDG 1 DGDG 2 PG	
PsbC (CP43)	450 (12-461)	13 Chl *a*	4 BCR	3 DGDG 3 PG	
PsbD (D2)	341 (12-352)	2 Chl *a*	1 BCR	3 PG	1 plastoquinone 1 BCT
PsbE	68 (13-80)				
PsbF	30 (15-44)				1 haem
PsbH	64 (20-83)		1 BCR	1 DGDG	
PsbI	34 (1-34)				
PsbJ	31 (8-38)				
PsbK	36 (11-46)				
PsbL	37 (2-38)			1 PG	
PsbM	32 (2-33)				
PsbO	210 (52-209, 240-291)				
PsbTc	30 (1-30)		1 BCR		
PsbW	54 (60-113)				
PsbX	35 (64-98)				
PsbZ	62 (1-62)			1 MGDG	
Psb30	30 (4-33)				
CP26	236 (47-282)	9 Chl *a* 4 Chl *b*	2 Lut 1 Neo	1 PG	
CP29	223 (11-233)	10 Chl *a* 3 Chl *b*	1 Lut 1 Vio 1 Neo	1 PG	
S-LHCII (Trimer)	657 (30-248)	24 Chl *a* 18 Chl *b*	6 Lut 3 Vio 3 Neo	3 PG	
M-LHCII (Trimer)	657 (30-248)	24 Chl *a* 18 Chl *b*	6 Lut 3 Vio 3 Neo	3 PG	
N-LHCII (Trimer)	657 (30-248)	24 Chl *a* 18 Chl *b*	6 Lut 3 Vio 3 Neo	3 PG	
Total	4810	189	53	34	6

Abbreviations used: BCR, β-carotene; Lut, lutein; Neo, neoxanthin; Vio, violaxanthin; MGDG, monogalactosyldiacyl glycerol; SQDG, sulfoquinovosyldiacyl glycerol; DGDG, digalactosyldiacyl glycerol; PG, phosphatidyl glycerol; BCT, bicarbonate ion.