



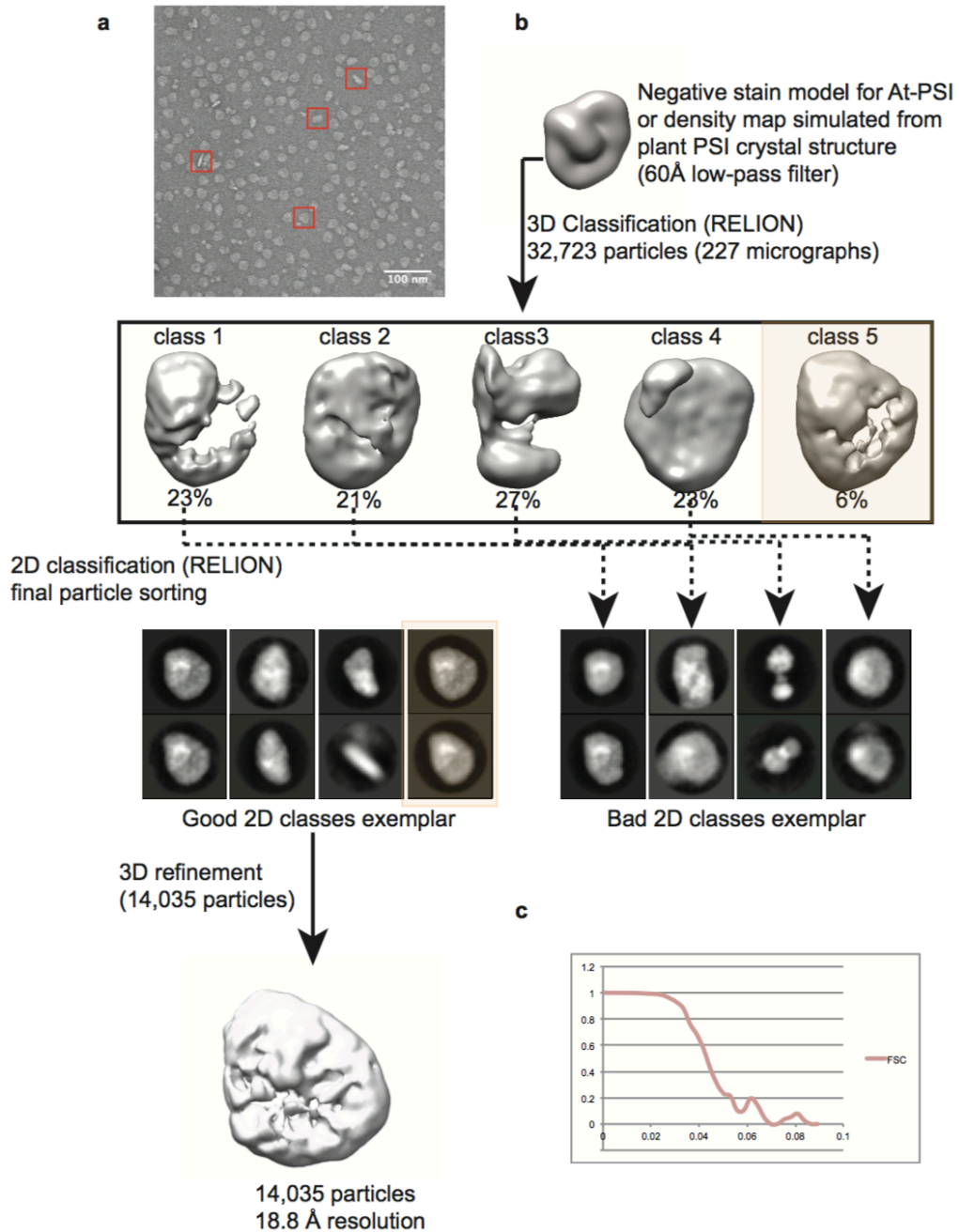
In the format provided by the authors and unedited.

A unique supramolecular organization of photosystem I in the moss *Physcomitrella patens*

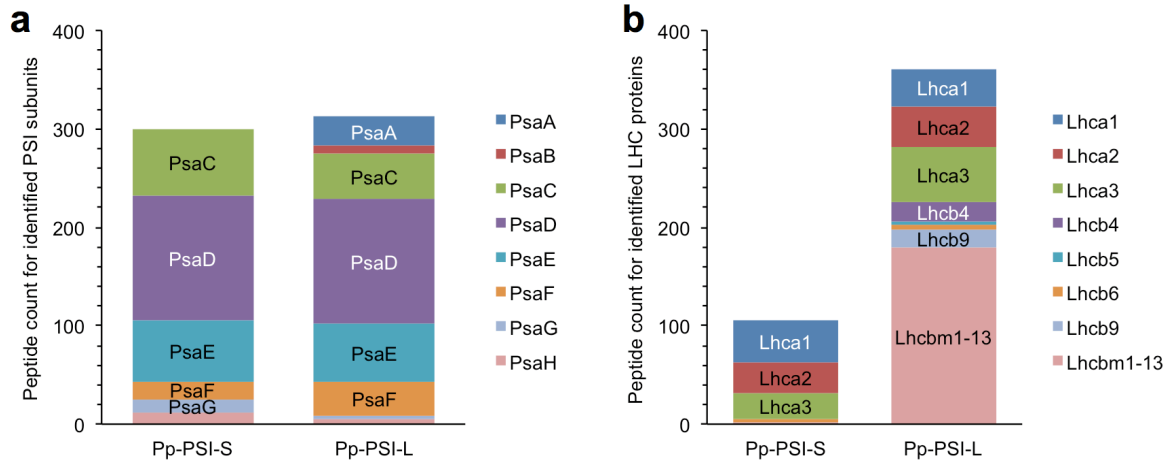
Masakazu Iwai ^{1,2,6*}, Patricia Grob^{2,3,6}, Anthony T. Iavarone⁴, Eva Nogales^{2,3,5} and Krishna K. Niyogi ^{1,2,3*}

¹Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA. ²Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ³Howard Hughes Medical Institute, University of California, Berkeley, CA, USA.

⁴QB3/Chemistry Mass Spectrometry Facility, University of California, Berkeley, CA, USA. ⁵Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. ⁶These authors contributed equally: M. Iwai, P. Grob. *e-mail: miwai@berkeley.edu; niyogi@berkeley.edu

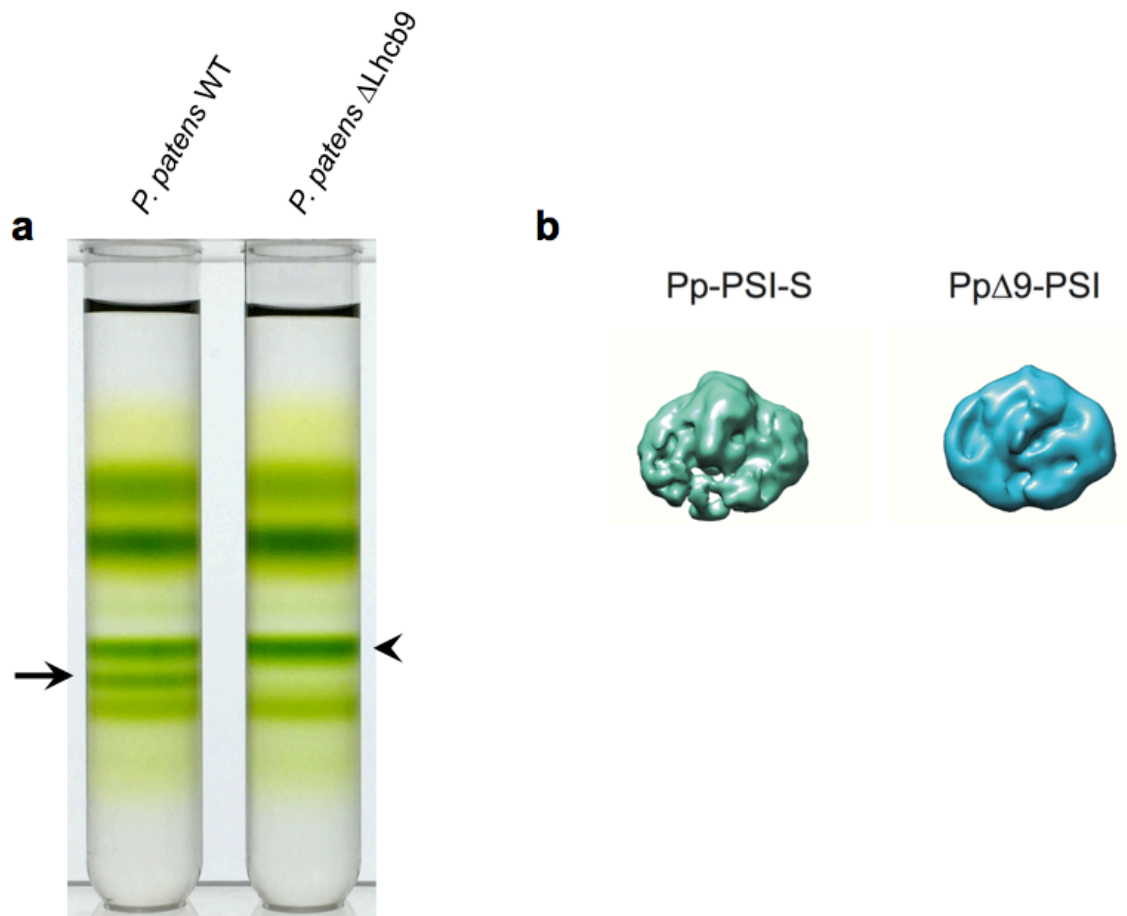


Supplementary Figure 1 | Image processing scheme in RELION for the negative stain EM data. Cr-PSI was used as example. **a**, A representative negative stain micrograph (from 227 micrographs) with a few particles boxed out. The data set size is summarized in Supplementary Table 2. **b**, Data was sorted in 3D, starting with the low-pass filtered (60 Å) plant PSI model as initial reference (At-PSI or *P. sativum* PDB 5L8R). Each 3D class was further analyzed by 2D classification and visual inspection. Classes corresponding to other complexes (ATP-synthase and PSII super-complexes in this example), aggregates, low quality and partial complexes were rejected while “good” classes were selected for 3D refinement. **c**, Final resolution was determined using a Fourier shell correlation (FSC) cutoff of 0.143.

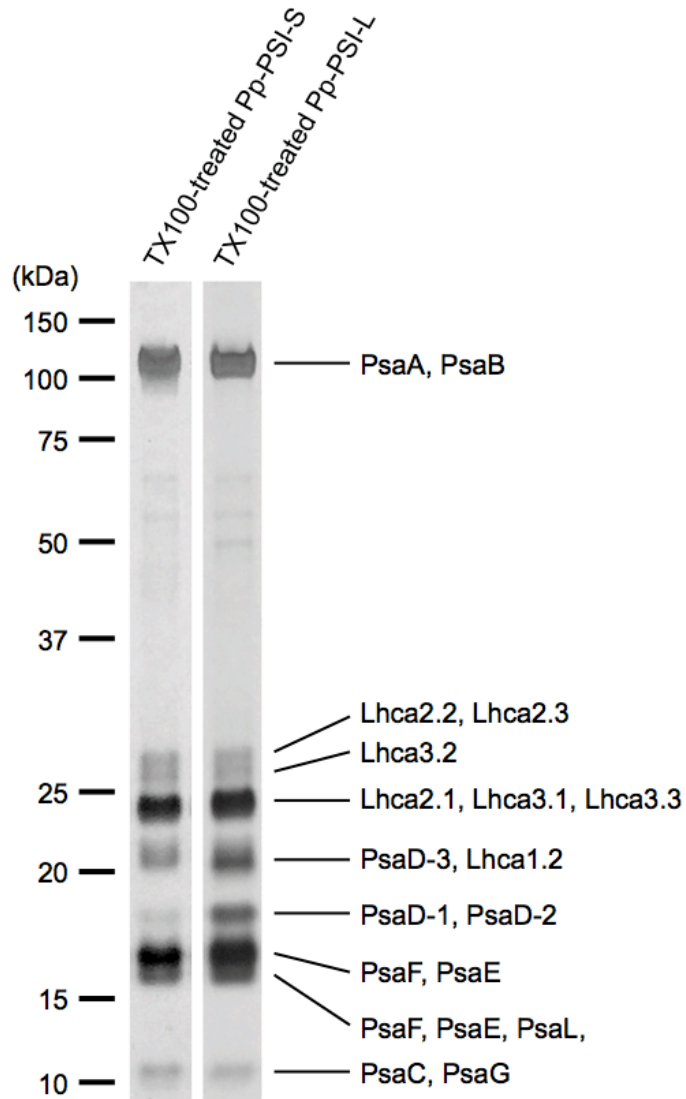


Supplementary Figure 2 | Protein compositions in Pp-PSI-S and Pp-PSI-L identified by MS.

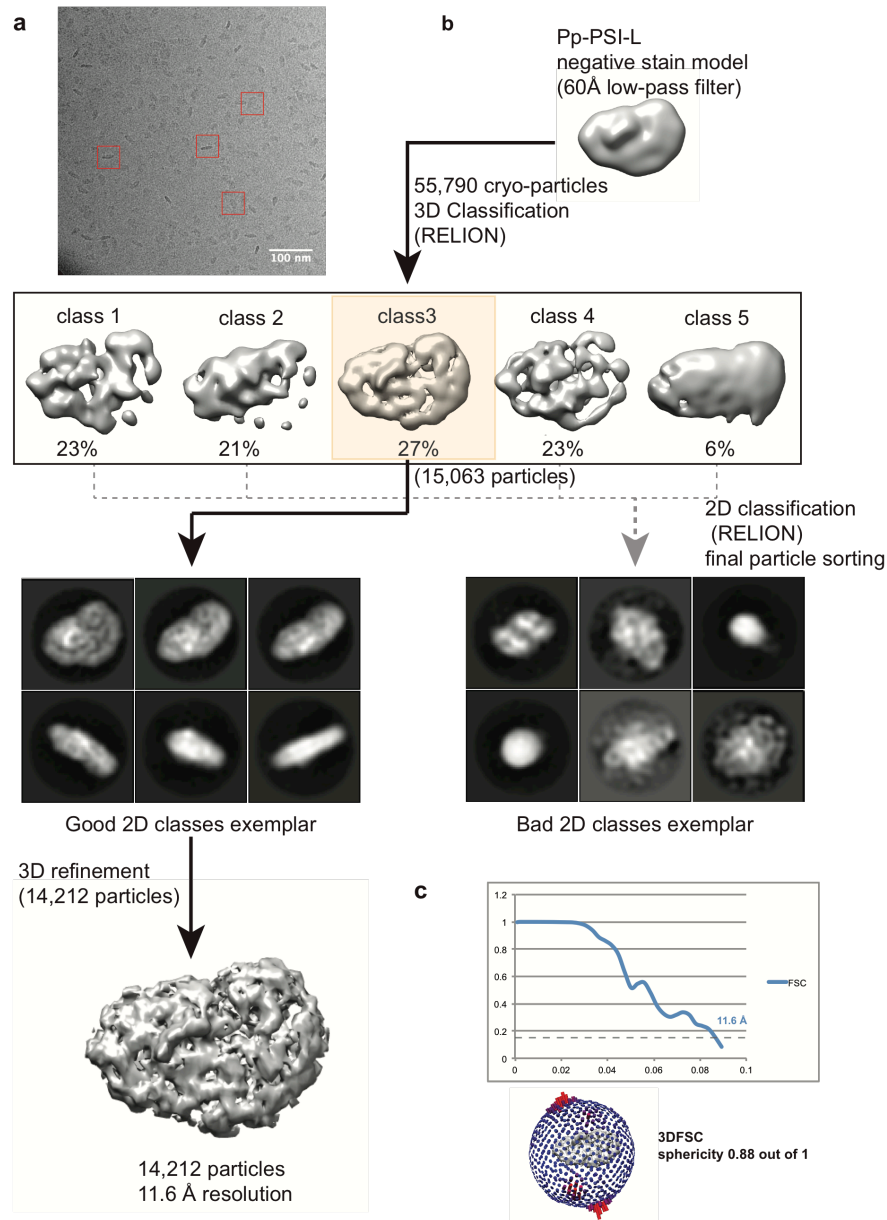
The Pp-PSI-S and Pp-PSI-L were collected separately, and in-solution digestion was done for MS analysis. **a**, Peptide count identified as the PSI subunits. **b**, Peptide count identified as LHC proteins. Note that it is possible that all digested peptides were not detected due to unknown factors involving efficiency of trypsin digestion of proteins and some limitations for mass spectrometry measurements such as different levels of ionization. See Supplementary Table 3 for details.



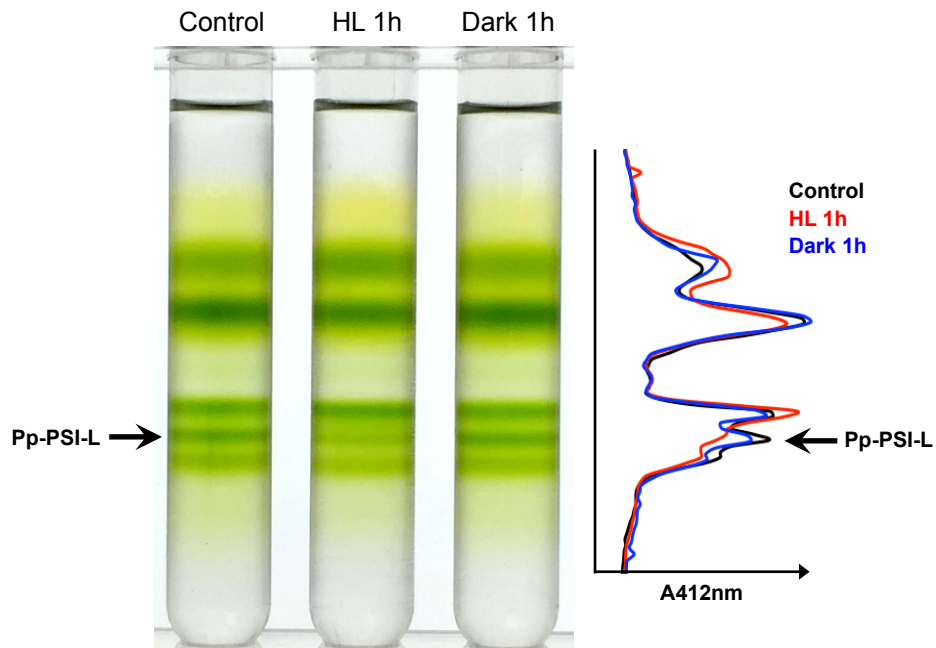
Supplementary Figure 3 | Lack of Pp-PSI-L in *P. patens* Δ Lhcb9 line. **a**, Maltose density gradient ultracentrifugation to isolate photosystem supercomplexes. The arrow indicates the position of Pp-PSI-L in WT. A representative result was shown from three technical replicates. **b**, The 3D reconstruction of the fraction containing PSI-LHCI supercomplex in Δ Lhcb9 (Pp Δ 9-PSI), which is indicated by the arrowhead in **a**. As a comparison, Pp-PSI-S from WT is also shown.



Supplementary Figure 4 | Protein composition of the distinct green bands from Pp-PSI-S and Pp-PSI-L after TX100 treatment. The distinct green bands as in Fig. 2a were separated by SDS-PAGE and stained by silver. The indicated bands were excised and digested by trypsin. The indicated proteins were identified by MS analysis. Representative results were shown from two technical replicates. See also Supplementary Table 4 for details.

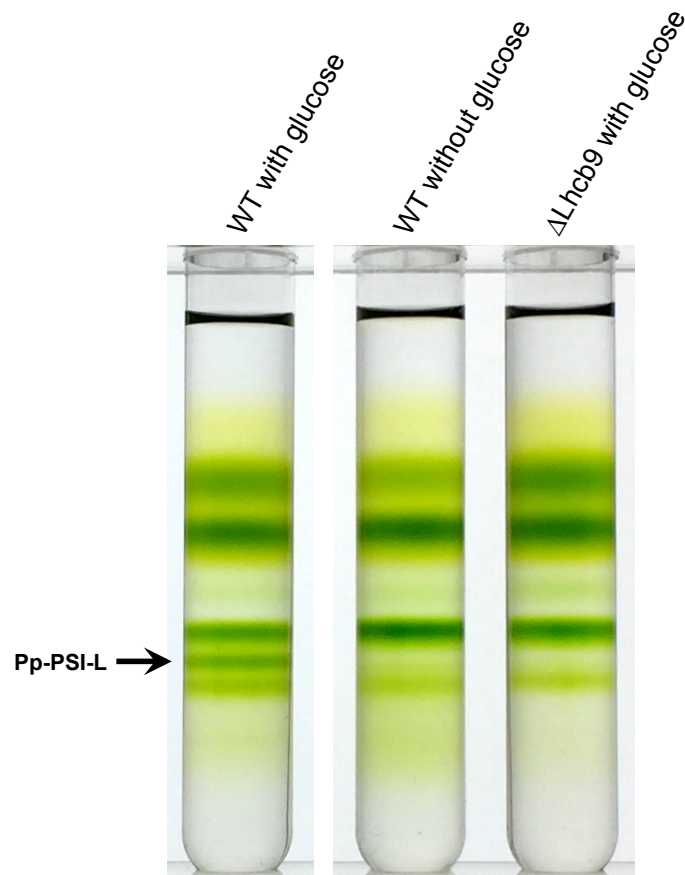


Supplementary Figure 5 | Classification and refinement scheme for the Pp-PSI-L cryo-EM structure. **a**, A representative cryo-EM image of the Pp-PSI-L complex (from 361 micrographs) with a few representative particles boxed out. The data set size is summarized in Supplementary Table 2. **b**, The processing scheme was similar to the one used for the negative stain data, with the exception of the initial model, which in this case comes from the negative stain Pp-PSI-L structure. Five 3D classes were used to sort out the data and a final 25% of the initial 55,790 particle data set contributed to the 11.6-Å resolution 3D reconstruction obtained after refinement. **c**, Final FSC plot and angular distribution.

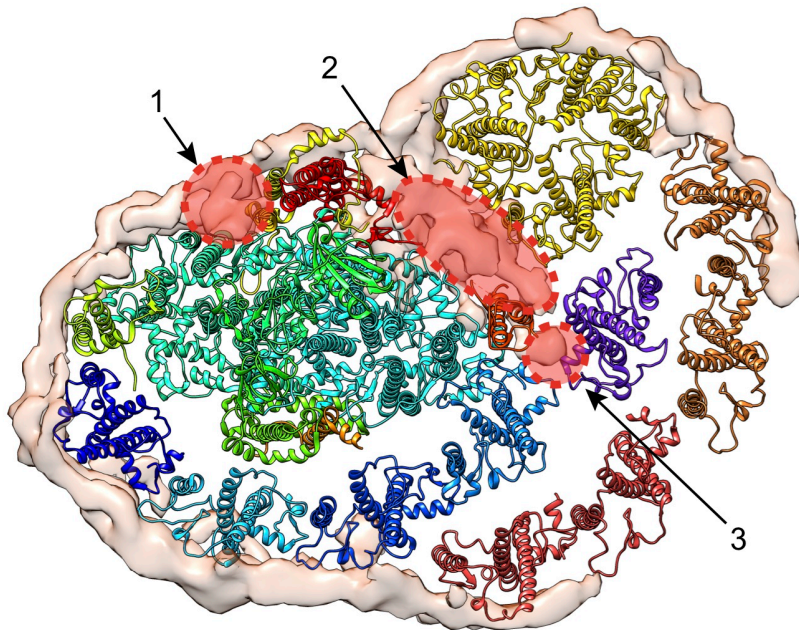


Supplementary Figure 6 | Reduced presence of Pp-PSI-L following high-light treatment.

The thylakoid membranes were isolated after either high light (HL; $1000 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ for 1 h) or dark treatment for 1 h, and samples were analyzed by maltose density gradient ultracentrifugation as described in the main text. The intensity of the band corresponding to Pp-PSI-L (indicated by an arrow) became lower after HL treatment. The absorbance (A) at 412 nm of each fraction is also indicated. A representative result was shown from two technical replicates.



Supplementary Figure 7 | Lack of Pp-PSI-L under growth conditions lacking an external carbon source. *P. patens* WT protonema cells were grown on BCDAT agarose media with or without glucose (5 g/L), under continuous light for 3 days. The thylakoid membranes were isolated and analyzed by maltose density gradient ultracentrifugation as described in the main text. The band corresponding to Pp-PSI-L (indicated by an arrow) was completely missing in the WT grown without glucose. The result of Δ Lhcb9 line grown with glucose is also shown for comparison. Representative results were shown from three technical replicates.



Supplementary Figure 8 | The difference map showing the unassigned density in

Pp-PSI-L. We used Chimera molmap to measure the size of unassigned density and compare it to putative components of the Pp-PSI-L. After excluding densities at the periphery, which are likely to be detergent or remaining lipids bound to the complex, we were left with mainly 3 small densities (dotted red circles). Since some of the unassigned density is probably due to lipids, detergents, and bound pigments, which are not discernable at this resolution, we did not quantify them as percentage of the mass. Instead, we compared their size to known proteins modeled with Chimera molmap. The circle 1 indicates a possible transmembrane density near PsaI that is large enough for a single helix and could correspond to one of the smallest PSI subunits, such as PsaM, but this has yet to be corroborated. The circle 2 indicates the largest density between PsaL, PsaA, and LHCII trimer that is large enough to have two trans-membrane helices and can tentatively be assigned to a small PSI subunit such as PsaO, as was determined for the recently published red alga PSI-LHCI and maize PSI-LHCI-LHCII complex structures (Pi X. et al., PNAS, 2018; PMID: 29632169; Pan et al., Science, 2018; PMID: 29880686). The circle 3 indicates a small nearby density that is only large enough for the PsaK N-terminus. Confirming this putative assignment for these extra densities and any further modeling will require a higher resolution map.

Supplementary Table 1 | Peptide counts identified in the lighter fraction produced from the TRX-treated Pp-PSI-L in the second run of maltose density gradient ultracentrifugation. MS analysis revealed the protein compositions in the lighter fraction produced from the TRX-treated Pp-PSI-L, which was originated from the additional LHC complex as shown in Figure 2a.

LHC proteins	Accession	Peptide counts
Lhca1.1	A9T399	7
Lhca1.2	A9T3X6	6
Lhca1.3	A9TJX6	6
Lhca2.1	A9RW10	9
Lhca2.2	A9TV00	17
Lhca3.1	A9TEM8	14
Lhca3.3	A9TB65	10
Lhcb3	A9TKU6	14
Lhcb4.1	A9U3M1	32
Lhcb4.2	A9T2F8	43
Lhcb5	A9U5B1	6
Lhcb6	A9RU32	3
Lhcb9.1	A9RR33	18
Lhcbm1	A9RNP2	24
Lhcbm2	A9RT62	35
Lhcbm4	Q765P5	22
Lhcbm7	A9T914	38
Lhcbm9	A9SGM1	13
Lhcbm10	Q765P6	15
Lhcbm13	A9RJS3	23

Supplementary Table 2 | Summary of all EM data. Total data collected, data selected through classification in RELION and the final 3D reconstruction resolution after refinement determined by Fourier shell correlation at the 0.143 cutoff after refinement when applicable.

PSI-LHCI supercomplexes	Number of micrographs	Total number of particles in dataset	Final number of particles after sorting	% of population	Number of 3D classes	Final resolution (Å)
At-PSI	213	27900	18499	66	5	16.8
Cr-PSI	227	32800	14035	43	5→4→3	18.9
Pp-PSI-S	392	40861	9266	23	5	16.7
Pp-PSI-L	619	56635	12700	22	3→2	21.1
PpΔb9-PSI	316	18873	8225	43	3	19.9
Pp-PSI-S-TX100	90	14480				
Pp-PSI-L-TX100	180	30846				
Pp-PSI-L cryo-EM	361	55780	14412	26	5	11.6

Supplementary Table 3 | Peptide counts identified in Pp-PSI-S and Pp-PSI-L by MS. The Pp-PSI-S and Pp-PSI-L were collected separately, and in-solution digestion was done for MS analysis. See also Supplementary Figure 2.

PSI subunits	Accession	Pp-PSI-S	Pp-PSI-L
PsaA	Q8MFA3	0	29
PsaB	Q8MFA2	0	8
PsaC	Q6YXQ2	68	46
	A9REG3	72	60
PsaD	A9SGR0	54	44
	A9RBZ9	0	23
	A9SPD7	31	22
PsaE	A9TU20	17	25
	A9TVA9	15	12
	A9SHP6	11	13
PsaF	A9TWD4	7	8
	A9SUQ3	0	8
	A9RNK7	0	5
PsaG	A9SJ10	9	4
	A9SH79	4	0
PsaH	A9SL09	12	5
LHC proteins	Accession	Pp-PSI-S	Pp-PSI-L
Lhca1.1	A9T399	15	11
Lhca1.2	A9T3X6	14	11
Lhca1.3	A9TJX6	14	16
Lhca2.1	A9RW10	14	19
Lhca2.2	A9TV00	8	11
Lhca2.3	A9TJ06	4	12
Lhca2.4	A9U5E4	5	0
Lhca3.1	A9TEM7	22	35
Lhca3.3	A9TB65	4	20
Lhcb4.1	A9U3M1	0	11
Lhcb4.2	A9T2F8	0	9
Lhcb5	A9U5B1	0	4
Lhcb6	A9RU32	4	4
Lhcb9.1	A9RR33	0	19
Lhcbm1	A9RNP2	0	13
Lhcbm2	A9RT62	0	23
Lhcbm4	Q765P5	0	18
Lhcbm5	A9S6S7	0	22
Lhcbm6	A9TL35	0	21
Lhcbm7	A9T914	0	28
Lhcbm8	A9RBQ3	2	0
Lhcbm10	Q765P6	0	32
Lhcbm13	A9RJS3	0	22

Other proteins	Accession	Pp-PSI-S	Pp-PSI-L
ATP synthase α subunit	Q6YXK3	62	25
	Q1XGA4	0	12
	P80658	78	44
ATP synthase β subunit	A9RKV9	23	5
	A9T949	17	4
	A9T278	11	8
ATP synthase γ subunit	A9TQE5	3	0
ATP synthase δ subunit	A9SYE0	5	0
ATP synthase ϵ subunit	Q6YXR2	15	4
D1 (PsbA)	Q6YXN7	0	5
D2 (PsbD)	Q6YXN8	0	8
CP43 (PsbC)	Q6YXN9	6	28
CP47 (PsbB)	Q6YXM8	10	23
	A9U1K9	25	24
PsbO	A9TSK0	19	19
	A9SA96	13	15
	A9RS28	12	0
PsbP	A9SG93	6	0
	A9SY12	5	0
PsbE (Cyt b559)	Q6YXL9	3	6
PsbR	A9S7M9	2	2
	A9RLG5	0	4
Psb32	A9SV74	0	2

Supplementary Table 4 | Protein composition of the distinct green bands from Pp-PSI-L after TX100 treatment. The distinct green bands as in Figure 2a were separated by SDS-PAGE and stained by silver (Supplementary Figure 4). The indicated bands were excised and digested by trypsin. The digested peptides were analyzed by MS.

PSI subunits	Accession	Peptide counts
PsaA	Q8MFA3	15
PsaB	Q8MFA2	11
PsaC	Q6YXQ2	9
	A9REG3	22
PsaD	A9SGR0	27
	A9RBZ9	6
	A9SPD7	16
PsaE	A9TVA9	4
	A9SHP6	13
PsaF	A9TWD4	27
	A9RNK7	11
	A9SJ10	3
PsaG	A9SH79	1
	A9RLG4	6
PsaL	A9S7M7	7
LHC proteins	Accession	Peptide counts
Lhca1.2	A9T3X6	9
Lhca2.1	A9RW10	9
Lhca2.2	A9TV00	8
Lhca2.3	A9TJ06	4
Lhca3.1	A9TEM7	23
Lhca3.2	A9U226	6
Lhca3.3	A9TB65	8