communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-026-09631-w

# Structural study of monomeric and dimeric photosystem I-LHCI supercomplexes from a bryophyte



Pi-Cheng Tsai [1], Romain La Rocca [1,2], Hiroyasu Motose [3], Jian-Ren Shen [1] ✉ & Fusamichi Akita [1] ✉

Photosystem I (PSI) is one of the two photosystems conserved from cyanobacteria to vascular plants, and associates with multiple light-harvesting complexes (LHCs) that capture and transfer solar energy. Liverworts such as *Marchantia polymorpha* occupy an early evolutionary position among land plants and faced major challenges during terrestrial adaptation, including desiccation, strong light, and UV radiation. We reveal the cryo-electron microscopic structures of PSI-LHCI monomer and homodimer from the liverwort *M. polymorpha* at resolutions of 1.94 and 2.52 Å, respectively. The high-resolution map allows identification of the cofactors of the monomer and reveal differences between the liverwort and moss, another clade of bryophytes. The PSI-LHCI monomer-monomer is stabilized by PsaG and PsaH interactions on the stromal side, which causes the bending and twisting of the homodimer. PsaM interacts with PsaB tightly, indicating a key role of PsaM in mediating the dimerization.

Photosystem I (PSI; EC 1.91.1.12) is one of two pigment-protein complexes responsible for light energy conversion in photosynthesis, and functions as a plastocyanin:ferredoxin oxidoreductase. Upon light excitation, the reaction center chlorophyll (Chl) P700 donates an electron to electron acceptor A0, and the electron is subsequently transferred through several acceptors and finally to ferredoxin at the stromal side[1,2]. On the other hand, the oxidized $P700^+$ is re-reduced by electrons provided by plastocyanin or cytochrome *c*6 at the lumenal side.

PSI is composed of a core complex and a peripheral antenna system. While most of the PSI core subunits are similar among cyanobacteria, various algae and vascular plants, the peripheral antenna system varies among different photosynthetic organisms[3–6]. In eukaryotes, PSIs contain varying numbers and structures of membrane-embedded light-harvesting complex I (LHCI) subunits, which bind different numbers of Chl and carotenoid pigments, and function to harvest the light energy and transfer it to the PSI core. Among them, vascular plants such as *Pisum sativum* bind four LHCIs at the PsaA/PsaB/PsaF/PsaJ side of the PSI core[7–9], whereas green algae such as *Chlamydomonas reinhardtii* contain 10 LHCIs, with eight bound to the PSI core as an inner and an outer LHCI belts at the same side as vascular plants, and the remaining two to the opposite side as an LHCI heterodimer (Supplementary Fig 1a, b)[10,11]. In bryophytes, the moss *Physcomitrium patens* has two types of PSI-LHCI supercomplex. A smaller form contains four LHCIs, which are similar to those of the vascular plants, whereas a larger form includes eight LHCIs arranged in two belts, which are connected by one LHCII trimer and one Lhcb9 (Supplementary Fig 1c)[12,13]. These structural variations may be important for the adaptation of various photosynthetic organisms to the different light and other environmental conditions that each organism experiences.

Different oligomerization states of the PSI core have been observed in different species. Either dimeric, trimeric, or tetrameric PSI cores are observed in cyanobacteria[14–17], and a tetramer PSI core has been observed in a glaucophyte *Cyanophora paradoxa*[18]. In land plants, PSI associated with its LHCI is purified and crystallized as a monomer initially[7–9]. However, there were reports showing that PSI-LHCIs from plants are observed as oligomeric structures by negative staining electron microscope (EM) and atomic force microscope (AFM), in which the dimeric structure is more abundant under dark[19,20]. Naschberger et. al. reported a PSI-LHCI dimeric structure from *C. reinhardtii* at 2.97 Å resolution by single particle cryo-electron microscopy (cryo-EM)[21]. The interface of the dimer is Lhca9-Lhca9. However, there are no LHCI proteins corresponding to Lhca9 in PSI of land plants, so there should be a different interface for PSI dimerization in plants, if the dimer indeed exists in plants.

*Marchantia polymorpha*, also known as a liverwort, is a non-vascular land plant and belongs to the bryophytes together with the mosses and the hornworts, and is one of the model species for investigating the structures and functions of land plants[22]. Two types of the PSI-LHCI supercomplex

[1]Research Institute for Interdisciplinary Science, Advanced Research Field, Okayama University, Okayama, Japan. [2]Institut de Biosciences et Biotechnologies d'Aix-Marseille (BIAM), Centre CEA Cadarache, Saint-Paul-lez-Durance, France. [3]Department of Biology, Graduate School of Environmental, Life, Natural Science and Technology, Okayama University, Okayama, Japan. ✉e-mail: shen@okayama-u.ac.jp; fusamichi_a@okayama-u.ac.jp

**Table 1 | Cryo-EM data collection, refinement and validation statistics for the *Mp*PSI-LHCI supercomplexes**

| Data collection and processing | PSI-LHCI monomer (EMDB-63404) (PDB 9LUT) | PSI-LHCI dimer (EMDB-63405) (PDB 9LUU) |
|---|---|---|
| Magnification | 165,000 | 165,000 |
| Voltage (kV) | 300 | 300 |
| Electron exposure ($e^{-}$ Å$^{-2}$) | 50 | 50 |
| Defocus range (μm) | −0.8 to −2.0 | −0.6 to −1.8 |
| Pixel size (Å) | 0.727 | 0.727 |
| Symmetry imposed | C1 | C2 |
| Initial particle images (no.) | 1,078,038 | 2,140,369 |
| Final particle images (no.) | 157,648 | 25,750 |
| Map resolution (Å) | 1.94 | 2.52 |
| FSC threshold | 0.143 | 0.143 |
| Map resolution range (Å) | 1.94–2.8 | 2.52–3.9 |
| Refinement | | |
| Initial model used (PDB code) | 6L35 | This study |
| Model resolution (Å) | 2.13 | 2.94 |
| FSC threshold | 0.5 | 0.5 |
| Model resolution range (Å) | 2.13–2.26 | 3.04–3.20 |
| Model-sharpening B factor (Å$^2$) | −26.9 | −30.7 |
| Model composition | | |
| Non-hydrogen atoms | 36,238 | 69,774 |
| Protein residues | 3,217 | 6,430 |
| Ligands | 207 | 400 |
| Unknown Ligand | 0 | 4 |
| Water | 606 | 0 |
| *B* factors (Å$^2$) | | |
| Protein | 61.60 | 96.03 |
| Ligand | 59.05 | 90.74 |
| Water | 57.07 | - |
| R.m.s. deviations | | |
| Bond lengths (Å) | 0.007 | 0.008 |
| Bond angle (°) | 1.471 | 1.637 |
| Validation | | |
| MolProblity | 1.54 | 1.77 |
| Clash score | 5.98 | 7.70 |
| Rotamer outliers (%) | 1.35 | 1.50 |
| Ramachandran plot | | |
| Favored (%) | 97.42 | 96.66 |
| Allowed (%) | 2.49 | 3.34 |
| Disallowed (%) | 0.09 | 0.00 |

structures have been reported from the moss *P. patens* so far, which showed a monomeric PSI core surrounded by different numbers of LHCI subunits[12,13,23]. Like the moss, *M. polymorpha* is at the intermediate position between the aquatic green algae and land plants. Structures of PSI-LHCI supercomplex from *M. polymorpha* may provide more information for understanding the evolution of photosynthesis during the adaptation to terrestrial environments.

In this work, we determined the structure of a PSI-LHCI monomer binding 4 LHCIs and a PSI homodimer binding a total of 8 LHCIs at an overall resolution of 1.94 Å and 2.52 Å, respectively, from the liverwort *M. polymorpha*. The structure of the PSI-LHCI monomer is very similar to that of the moss *P. patens*. The native PSI-LHCI dimer is isolated from a mixture of PSI-LHCI monomer and dimer, which has a monomer-monomer interface involving PsaM, PsaB, PsaI, PsaG and PsaH. We reveal the structure of the PSI-LHCI dimer, which suggests that the PsaM and PsaB interaction plays an important role in dimerization of PSI-LHCI in bryophytes.

## Results

### Overall structure of PSI-LHCI from *M. polymorpha*

The PSI-LHCI supercomplexes were purified as described in the Methods section. Both monomers and dimers were obtained, with dimers obtained at a lower concentration of n-dodecyl-α-D-maltoside (α-DDM) than that used for monomers, suggesting a less stable conformation of the dimer than the monomer. Analyses with clear native-PAGE, SDS-PAGE, absorption spectrum and negative-staining electron microscopy showed that both purified PSI-LHCI monomer and dimer contained a PSI core and LHCI subunits (Supplementary Fig 2 and Supplementary Data 2). The density map of the monomeric PSI-LHCI supercomplex was reconstructed from 157,648 particles at a resolution of 1.94 Å, from which most of the amino acid side chains, water molecules and light-harvesting pigments in the PSI core and LHCIs could be built based on the clear density map (Supplementary Figs. 3–5 and Table 1). In total, 606 water molecules are assigned in the PSI-LHCI monomer structure, most of which are at the surface of the PSI subunits, except 11 water molecules near 2 phylloquinone (PQNs) and the water molecules coordinated to Chl.

The density map for the dimeric PSI-LHCI supercomplex was reconstructed from 25,750 particles at a resolution of 2.52 Å, and most of the amino acid side-chains and light-harvesting pigments could be built based on the density map (Supplementary Figs. 3–5 and Table 1).

*M. polymorpha* PSI-LHCI (hereafter denoted as *Mp*PSI-LHCI) monomer is composed of 13 subunits in the core complex (PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaG, PsaH, PsaI, PsaJ, PsaK, PsaL, PsaM) and 4 LHCIs (Lhca1, 2, 3, 4), which are bound at the PsaK/PsaF/PsaJ/PsaG side of the core (Fig. 1a, b). The overall structure of monomeric *Mp*PSI-LHCI is similar to that of *P. patens*[12,13,23], although the loops of PsaL and PsaH on the stromal side were not visible in the present structure. The model of dimeric *Mp*PSI-LHCI is composed of two sets of all 17 subunits in the core and LHCIs, which are arranged in a rotational symmetric way (Fig. 1c, d). The monomer-monomer interface of the dimer is at the PsaH/PsaI/PsaM/PsaB/PsaG side, where a few signals between the two monomers were observed, but they cannot be identified due to their weak intensities. On the other hand, the invisible loops of PsaL and PsaH in the monomer are visible in the dimeric structure.

The components of the electron-transfer chain are assigned unambiguously, which consists of a special pair of Chls P700, accessory Chl $A_{-1}$, primary electron acceptor Chl $A_0$, secondary electron acceptor $A_1$ (phylloquinone), and three iron-sulfur clusters $F_X/F_A/F_B$ (Fig. 2a and Supplementary Fig 5). Their arrangement is shown in Fig. 2a, where Chl *a* and the $C13^2$ epimer form of Chl *a*, designated Chl *a'* that constitute the pair of P700 can be distinguished clearly from the high-resolution map (Supplementary Fig 5b, c). There are slight differences among the electron transfer cofactors between the present structure and the previously reported *P. patens* structure[12,13,23]. The distance between PQN ($A_1$) and the closest iron atom of $F_x$ in the A branch is 9.11 Å in *M. polymorpha*, which is 0.17 Å longer than that in *P. patens* (Fig. 2a). In contrast, the distance between PQN and $F_x$ in the B branch is 9.00 Å, which is 0.14 Å shorter than that in *P. patens* (Fig. 2a). The distance between $F_x$ and $F_A$ is 12.53 Å in *M. polymorpha*, which is 0.17 Å longer than that in *P. patens*, and the distance between $F_A$ and $F_B$ is 9.43 Å, which is 0.14 Å longer than that in *P. patens* (Fig. 2a).

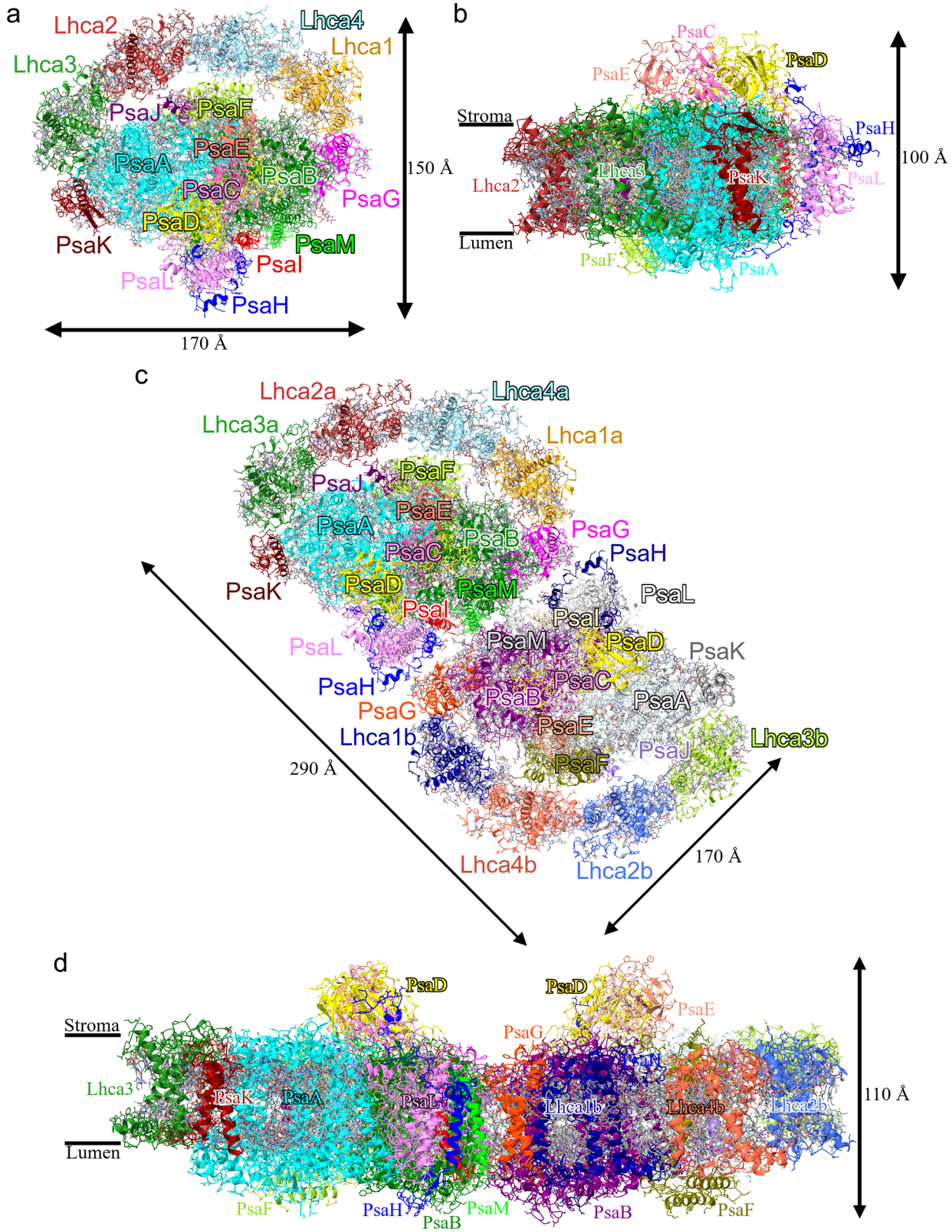


**Fig. 1 | Overall structures of the *Mp*PSI-LHCIs monomer and dimer supercomplexes. a** Overall structure of the *Mp*PSI-LHCIs monomer viewed from the stromal side. **b** Structure of the *Mp*PSI-LHCIs monomer with a view along the membrane plane. **c** Overall structure of the *Mp*PSI-LHCIs dimer viewed from the stromal side. **d** Structure of the *Mp*PSI-LHCIs dimer with a view along the membrane plane. Color codes: PsaA, cyan and light cyan; PsaB, green and purple; PsaC, hot pink and pink; PsaD, yellow and gold; PsaE, salmon and dark salmon; PsaF, green yellow and oliver; PsaG, magenta and orange red; PsaH, blue and navy; PsaI, red and wheat; PsaJ, purple and light purple; PsaK, dark red and gray; PsaL, violet and white; PsaM, lime and light gray; Lhca1, orange and dark blue; Lhca2, brown and royal blue; Lhca3, forest green and yellow green; Lhca4, sky blue and tomato.

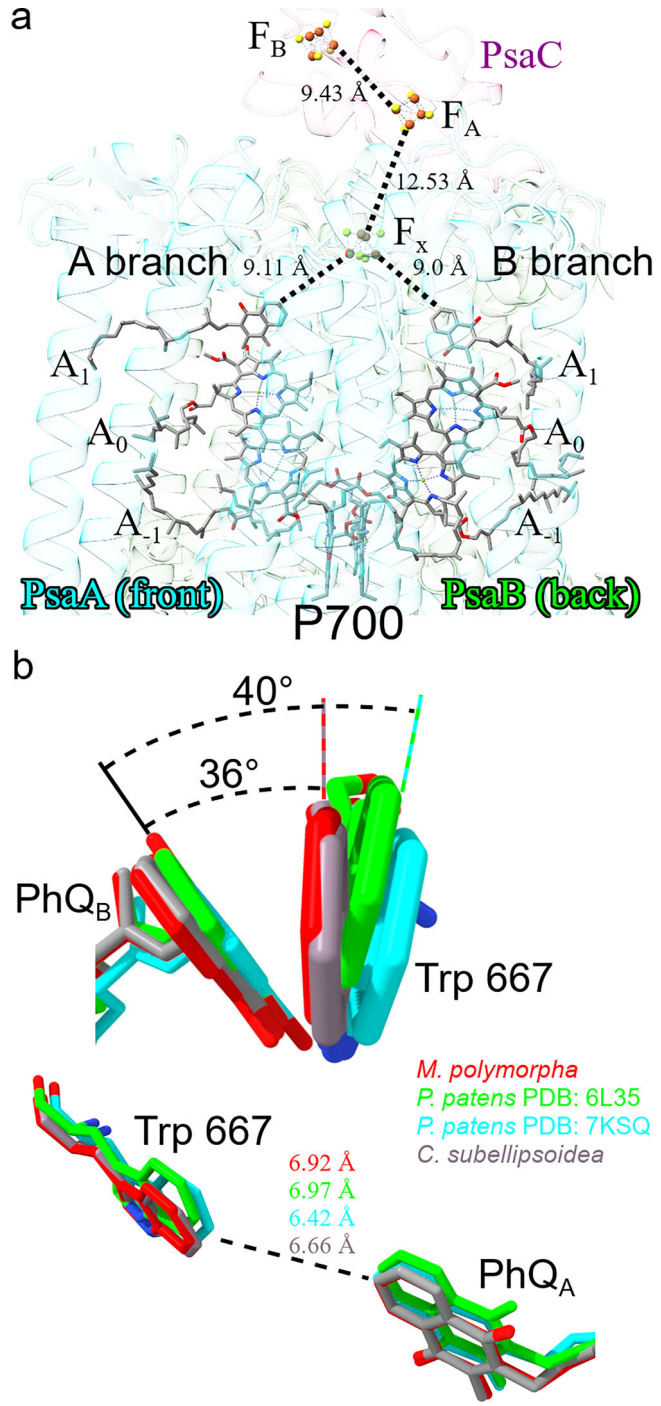


**Fig. 2 | Electron-transfer chain of the *Mp*PSI-LHCIs supercomplex.** **a** Arrangement of cofactors involved in the electron-transfer chain. P700, special pair Chls; $A_{-1}$, accessory Chl; $A_0$, primary electron acceptor; $A_1$, secondary electron acceptor phylloquinone; $F_X$, $F_A$, and $F_B$, iron-sulfur clusters. The distances between cofactors are measured between the edge-to-edge of the closest atoms in each cofactor. **b** The angle between the aromatic ring planes of Trp667 and the phylloquinone of the B branch, and the distance between Trp667 and the phylloquinone of the A branch. Color codes: *M. polymorpha*, red; *P. patens* PDB:6L35, lime; *P. patens* PDB:7KSQ, cyan; *C. subellipsoidea*, gray.

However, these differences are small and may fall into the error range of the experimental data.

Another difference in the cofactor environment is found at the Trp667 residue in PsaB. In the structure of PSI from a thermophilic cyanobacterium *Thermosynechococcus elongatus* determined by cryogenic crystallography [PDB: 1JB0][17] and room temperature by X-ray free electron laser [PDB: 7M75][24], the angle between the aromatic ring planes of Trp and $PhQ_B$ is changed from 35° to 50° and the distance between the Trp and the $PhQ_A$ is changed from 6.6 to 7 Å, from cryogenic temperature to room temperature (Supplementary Fig 6). In our structure, the angle is 36° which is similar to the cryogenic structure, whereas this angle is around 40° in the PSI structure of *P. patens* which is also determined by cryo-EM [PDB: 6L35 and 7KSQ] (Fig. 2b, top)[12,13,23]. On the other hand, the distance between the Trp and $PhQ_A$ is 6.9 Å in our structure as well as the structure of the *P. patens* [PDB: 7KSQ][23] (Fig. 2b, bottom), but it is 6.4 Å in another structure of *P. patens* [PDB: 6L35][12]. In the structure of PSI-LHCI from an Antarctica green alga *Coccomyxa subellipsoidea* determined at a resolution of 1.92 Å [PDB: 9KQP], this distance was shown to be 6.7 Å[25]. As the previous structures are solved at resolutions ranging from 2.8 to 3.2 Å, we consider that these differences may be due to different resolutions, and our distance and angle reflect the real situation in a more accurate way than before. The angle of 36° between the ring planes of Trp and $PhQ_B$ in the current structure is similar to the distance in the cryogenic structure of *T. elongatus* PSI, but the distance of 6.9 Å between the trypsin and the $PhQ_A$ is similar to the room temperature structure of *T. elongatus* PSI (Fig. 2b and Supplementary Fig 6). Thus, the differences in the angle and distance between PhQ and Trp may be due to factors other than temperature, and it is possible that these differences may be due to errors in the experimental data.

Because of the high resolution of the map, we can identify the water molecules that coordinate Chls (Supplementary Table 1 and 2). There are 33 out of 154 Chls coordinated by water molecules, among which 17 are in LHCIs and 16 are in the PSI core. PsaK/CLA203 is not visible in the dimer because of the poor density map. Förster resonance energy transfer (FRET) rate constants ($K_{FRET}$) between adjacent Chls are calculated, and potential FRET pathways with FRET rates higher than 0.1 $ps^{-1}$ were illustrated by dashed lines connecting the Mg atoms in Fig. 3, and their rates are shown in Supplementary Data 1, which showed no remarkable differences between *Mp*PSI-LHCI monomer and dimer. This indicates that the interface between the two monomers does not feature any pair of Chls at a suitable distance and orientation for an energy transfer.

### Arrangement of light-harvesting complex subunits and pigments in LHCI

The structures of 4 LHCIs from *M. polymorpha* are superposed, which showed that helices A to D have no remarkable differences among the 4 LHCIs. However, conformational changes exist in the loop regions. The BC loop of Lhca1 and Lhca3, located in the lumenal side, is longer than the other two LHCIs (Fig. 4a, boxed area with dashed lines), because helix B of Lhca1 ends earlier and thus is around 3 amino acids shorter than the other 3 LHCIs in addition to the length change of amino acid sequences (Supplementary Fig 7). Because the helix B of Lhca3 starts later and the amino acid sequences are longer in the AC loop (because this loop is long, it is divided into AC-A and AC-C loops that are close to A-helix and C-helix, respectively, in Fig. 4a) and BC loop of Lhca3, the N-terminus, AC loop and BC loop of Lhca3 are longer than other LHCIs (Fig. 4a, boxed area with dotted lines and dash-dotted lines), in which, the longer BC loop of Lhca3 provides an additional Chl-coordinated site (CLA313) (Fig. 4b, red arrow). The AC loop of Lhca4 close to helix A at the stromal side, is 3 amino acid residues shorter than the other 3 LHCIs (Fig. 4a, boxed area with solid lines). The amino acid sequence at this region is 4 or 5 residues shorter than other LHCIs, which is observed only in *M. polymorpha* but not in *P. patens*.

Most Chl-binding sites are similar in the 4 LHCIs, but multiple differences are also found (Fig. 4b, Supplementary Table 2). Chl *a*314/Lhca1 at the lumenal side is coordinated by the main chain of Ile237 at the C-terminal of Lhca1, and it exists only in Lhca1 of *M. polymorpha* and *P. patens* (Fig. 4b). In other Lhcas, the isoleucine is changed to leucine; however, this change should not disrupt the coordination of Chl *a* because the Chl *a* is coordinated by the main chain of the residue. Thus, there are other reasons for the difference in the presence/absence of this Chl among the 4 LHCIs. Chl *a*314/Lhca1 provides a potential FRET pathway to Chl *a*302/PsaF of the

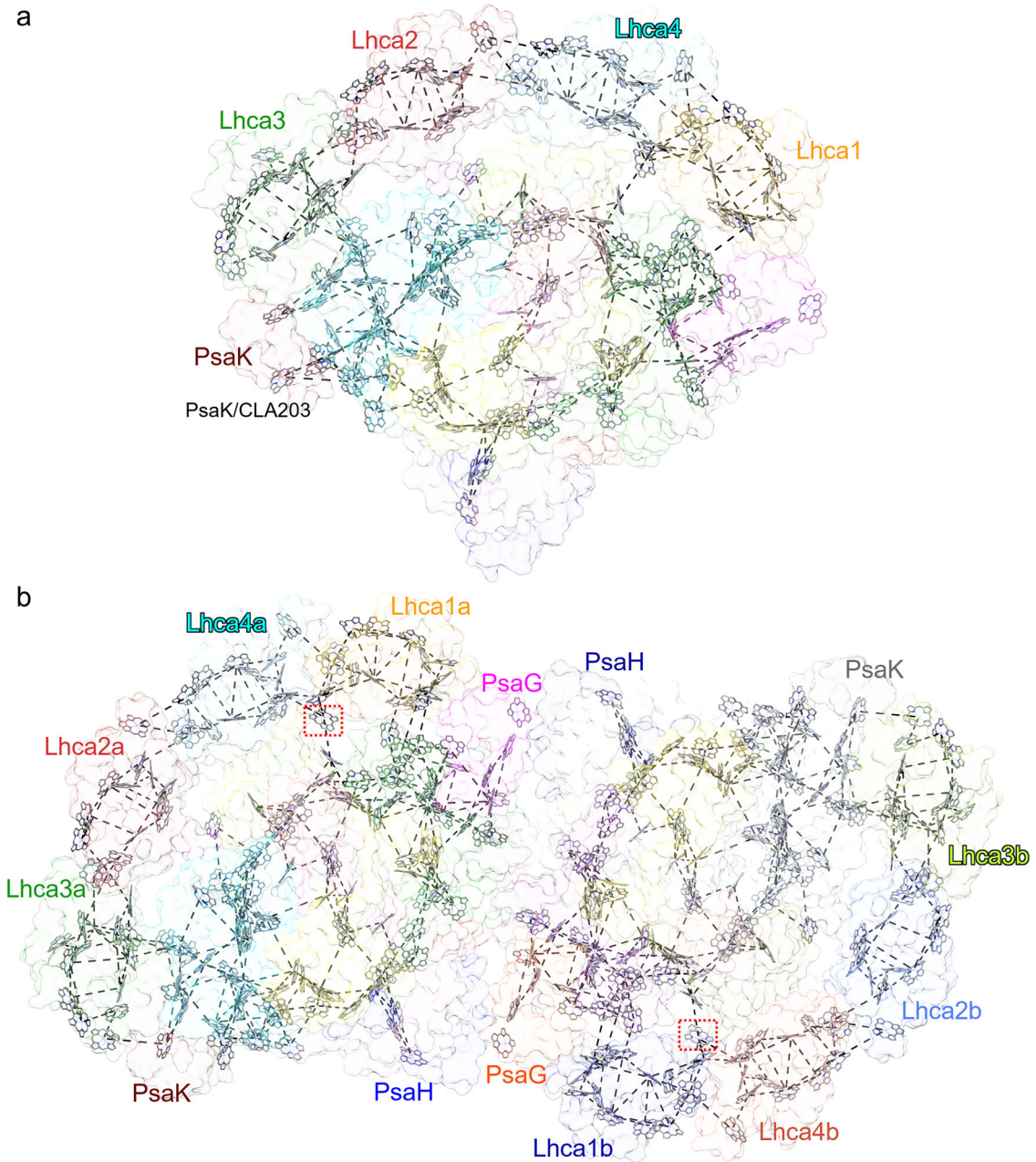


**Fig. 3 | Possible energy transfer pathways in the *Mp*PSI-LCHIs monomer and dimer supercomplexes.** View from the stromal side of *Mp*PSI-LHCI monomer (**a**) and dimer (**b**). The pathways between the Chls are connected between Mg of each Chls by dashed lines when the FRET rate is higher than 0.1 $ps^{-1}$. The color codes for each subunit are the same as those in Fig. 1.

PSI core, and it seems that this Chl is important for energy transfer from LHCIs to the core complex (Fig. 3, boxed area with red dotted lines). Chl *b*314/Lhca2 and Chl *b*314/Lhca4 at the stromal side are coordinated by aspartic acids of the AC loop, and this Chl does not exist in the other two Lhcas (Fig. 4b, orange arrow). Chl *a*313/Lhca3 at the lumenal side is coordinated by a water molecule near the BC loop (Fig. 4b, red arrow). The longer BC loop of Lhca1 occupied this space, and loops of Lhca2 and Lhca4 are shorter than the other two LHCIs; therefore, this Chl exists only in Lhca3. Chl *b*314/Lhca3 and Chl *a*315/Lhca4 are located in the middle of helix C of these subunits and coordinated by Lhca3-His174 and Lhca4-His148 residues in helix C, respectively (Fig. 4b, lime arrow). There is no histidine in helix C of Lhca1 and Lhca2, therefore, these 2 Chls are present only in Lhca3 and Lhca4, although the position of Lhca3-His174 and Lhca4-His148 are rather separated in the amino acid sequences of Lhca3 and Lhca4 (Supplementary Fig 7). Chl *a*305/Lhca3 is located at the lumenal side; however, this Chl is changed to Chl *b* in other three Lhcas (Chl *b*306/Lhca1, Chl *b*306/Lhca2 and Chl *b*306/Lhca4), and Chl *a*305/Lhca3 is coordinated by a valine residue of the longer BC loop, but this Chl is coordinated by water molecules in Lhca2 and Lhca4 and a glutamine residue in Lhca1 (Fig. 4b, blue arrow). Because of the longer BC loop in Lhca3, the position of Chl *a*305/Lhca3 is located farther from Lhca3 and becomes closer to Chl *a*817 of PsaA to establish a potential FRET pathway. Chl *a*310/Lhca1, Chl *a*310/Lhca2, Chl *a*309/Lhca3, and Chl *a*310/Lhca4 are counterparts at each Lhca with slight position shifts. In Lhca1 and Lhca2, this Chl is coordinated by a phospholipid (LHG); however, it is coordinated by a water molecule in Lhca3 and Lhca4, where there is either no LHG in Lhca3 or the LHG has shifted its position in Lhca4 (Fig. 4b, pink arrow). Regarding the position of other ligands, LUT318/Lhca1 near helix C has its counterparts in other

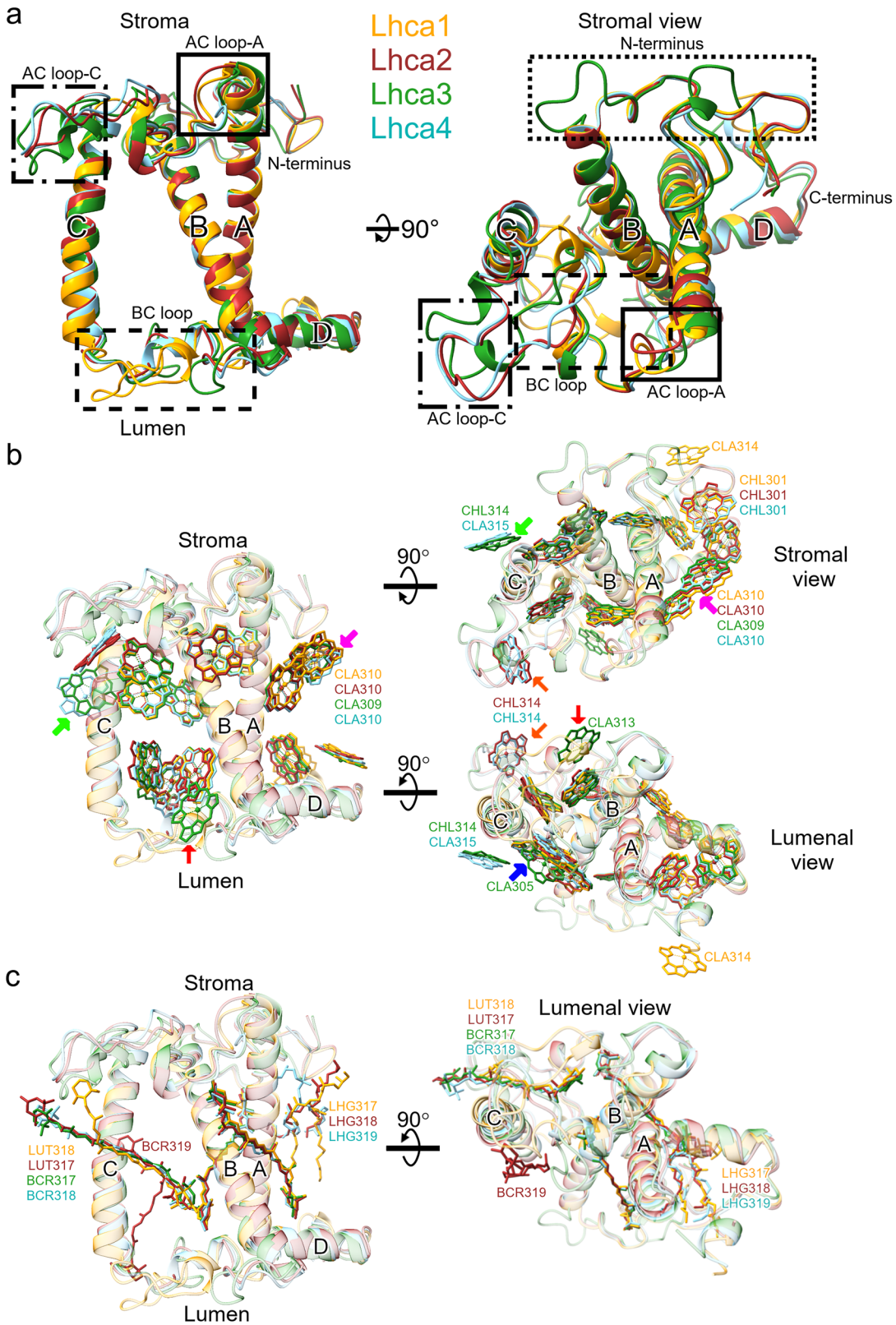


**Fig. 4 | Comparison of the arrangement of LHCIs, Chls and other ligands among 4 LHCIs. a** Overall arrangement of 4 LHCIs from a view along the membrane plane and the stromal side. The square with dotted lines shows the differences at the N-terminus, squares with solid lines and dash-dotted lines show the differences at the AC loop, and square with dashed-line shows the differences at the BC loop. **b** Overall arrangement of Chls among 4 LHCIs viewed along the membrane plane, lumenal side and stromal side, respectively. The number of Chls is shown only when it has differences among the 4 LHCIs. **c** Overall arrangement of other ligands among 4 LHCIs viewed along the membrane plane, lumenal side and stromal side, respectively. Only ligands that had differences among the 4 LHCIs are labeled together with their numbers. CLA: chlorophyll *a*; CHL; chlorophyll *b*. BCR, β-carotene; LHG, dipalmitoylphosphatidyl glycerol; LUT, lutein. The color codes are the same as in Fig. 1. The comparison is based on the superposition of 4 LHCIs structures in the *Mp*PSI-LHCIs monomer by MatchMaker in ChimeraX.

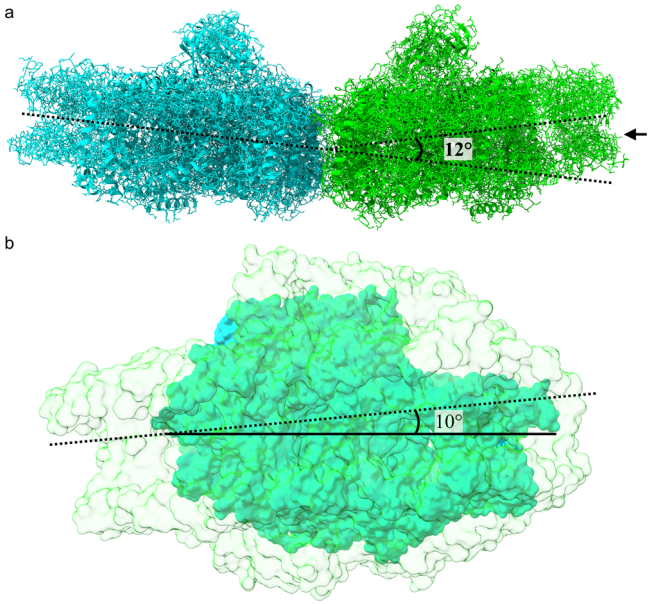


**Fig. 5 | Arrangement of the *Mp*PSI-LHCIs dimer. a** The angle between the two monomers in a dimer viewed from the longer side of the dimer along the membrane plane. There is around 12° bending between the two monomers. **b** The angle between the two monomers in a dimer viewed from the shorter side of the dimer along the membrane plane (black arrow). There is around 10° twisting between the two monomers.

Lhcas; however, one aromatic head of it bends to the stromal side remarkably in Lhca1. In addition, there is an extra BCR319 in Lhca2 near helix C (Fig. 4c), which does not exist in other Lhcas.

### Structural analysis of the dimerization interface of PSI-LHCI

In the PSI-LHCI dimer structure, the two monomers are not in the same horizontal plane. One monomer tilts around 12° towards the stromal side when observed from the longer side of the dimer along the membrane plane (Fig. 5a), and twists counterclockwise around 10° with another monomer when observed from the shorter side of the dimer along the membrane plane (black arrow in Fig. 5b). This may be due to stronger interactions at the stromal side than the lumenal side between the two monomers. Cryo-EM map was separately subjected to 3D variability analysis with cryoSPARC 4.6.0 (Supplementary Fig 8 and Supplementary Movie 1)[26], and the results show that the tilt and twist angle were not consistent, with part of the particles located in the same horizontal plane.

Because of the weak signals in the interface, this region may be accessible to the solvent. Nevertheless, the interactions between amino acid residues can still be determined. From the results of PISA analysis[27] which removes all interactions within each PSI-LHCI monomer, there are two regions identified as a linker between the monomers. One is the PsaM and PsaI region of one monomer, which interacts with PsaB from the adjacent monomer (Fig. 6a, boxed area with solid lines), and the other is the stromal region of PsaH, which may be partially assisted by PsaL, which interacts with PsaG and its associated co-factors from the adjacent monomer (Fig. 6a, boxed area with dotted lines). At the interface between PsaB, PsaM and PsaI, a total of 7 interacting areas are detected by PISA analysis (Fig. 6b and Supplementary Table 3). In area 1, the interface involves PsaB (I144, I148, L155, K160, W161, W209, D210, F212, L213, and T214) and PsaM (S5, D6, S7, I10, I18, R25, E29, and Q32) at a region from the stromal side to the lumenal side. In area 2, the interface involves two PsaB subunits (W154, L155, Q158, and W161) at the stromal side. In areas 3 and 4, the interface involves PsaM (S7, I10, V11, and L14) and the ligand of PsaB (Chl *a*813 and UNL854) at the stromal side. In areas 5, 6, and 7, the interfaces are between PsaB (L213, Chl *a*810, and BCR851), PsaM (R25) and PsaI (A3 and Y5) at the lumenal side. Most of the interacting amino acid residues are hydrophobic, which indicates that these interfaces are formed mainly by hydrophobic interactions. The interactions extend from the stromal side to the lumenal side, especially between PsaB and PsaM; therefore, PsaB and PsaM may play a key role in providing strong hydrophobic interactions in the binding of two PSI-LHCI monomers.

At the interface between PsaG and PsaH, a total of 4 interacting areas are detected by PISA analysis (Fig. 6c and Supplementary Table 3). In area 1, the interface is between Chl *a*202/PsaG and PsaH (E88, T89, W90, Y92, and P93). In area 2, the interface is between PsaG (S120, L121, K123, T124, and Q125) and PsaH (E88, T89, Y92, K96, and I99). In areas 3 and 4, T89, F90, and K102 of PsaH interact with BCR203 of PsaG and BCR853 of PsaB. All interactions between PsaG and PsaH are at the stromal side, which may be the reason for the tilt and twist between the two PSI-LHCI monomers.

## Discussion

The present study showed the structures of monomer and dimer of PSI-LHCI from a liverwort *M. polymorpha*. While the PSI monomer has been isolated from a plant by n-dodecyl-β-D-maltoside (β-DDM)[28,29], and its structure is solved by X-ray crystallography[7–9], we used α-DDM to solubilize the thylakoids to obtain both PSI-LHCI monomers and dimers in the present study. We do not know if it is due to the difference between β-DDM and α-DDM or due to species difference used in the two studies, but our PSI-LHCI dimer was obtained at a lower concentration of α-DDM than that used for the monomer, suggesting that the dimer may be more unstable than the monomer. On the other hand, PSI can form dimer, trimer and tetramer in cyanobacteria, and a dimer in the green alga *C. reinhardtii*. The observation of PSI dimer in *M. polymorpha* but not in land plants may suggest that the dimeric conformation is retained during evolution from green algae to bryophytes, but lost in land plants. As bryophytes live in a wet environment, this may further suggest that the formation of PSI dimer is not favored in land plants, which live in a land environment with sufficient light illumination. Thus, the PSI dimer may be a form adapted to a lower light environment.

The Chl binding sites in *M. polymorpha* were compared with the PSI-LHCI structure from *P. patens* (PDB: 7KSQ) and *Arabidopsis thaliana* (PDB: 8J7A) (Supplementary Fig 9). While most of the Chls are at a similar position in the three species, several Chl at the edge of PsaK/PsaA/PsaH were observed in *P. patens* and *A. thaliana* but they were not observed in *M. polymorpha* (left-lower boxed area with dotted lines of Supplementary Fig 9). These may be due to the reconstruction step of the PSI-LHCI map, instead of the absence of these Chls in PSI-LHCI in *M. polymorpha*. The other difference is between the LHCI-belt and PSI-core (boxed area with dotted lines of Supplementary Fig 9). The Chl(s) are observed in *P. patens* and *M. polymorpha*, not in the higher plant, *A. thaliana*. Because *P. patens* and *M. polymorpha* live under lower light conditions than *A. thaliana*, these Chls may provide another energy transfer pathway for higher efficiency. In *P. patens*, these two Chls provide a potential energy transfer pathway between LHCIs and PSI-core. However, in *M. polymorpha*, there is only one Chl at a similar position, and it seems that this Chl cannot provide a potential energy transfer pathway (Supplementary Fig 9).

The structure of Lhcas from *M. polymorpha* was compared with Lhcas from *C. reinhardtii*, *P. patens* and *A. thaliana*. According to the phylogenetic analysis, the comparison of structure and amino acid sequences was performed (Supplementary Fig 10). Lhca3 from 4 species was clustered together, and Lhca1 of 2 species was clustered together with Lhca6 of the other 2 species. Lhca2 of 3 plant species was clustered together, but no corresponding Lhca from *C. reinhardtii* was clustered with the land plants. No significant differences were found in the structures of 4 Lhca1 and 3 land plant Lhca2 (Supplementary Fig 11). On the other hand, Lhca3 has a longer N-terminus in *P. patens* and a longer C-terminus in *C. reinhardtii*. However, these regions were not included in the structure, so the Lhca3 structure is very similar between *P. patens* and *C. reinhardtii*. Lhca4 from *P. patens* has a longer AC loop than the other 3 species, which shows a conformation slightly different from the other 3 species (boxed area with red dotted lines of the bottom panel). Lhca4 of *M. polymorpha* has a very similar structure to

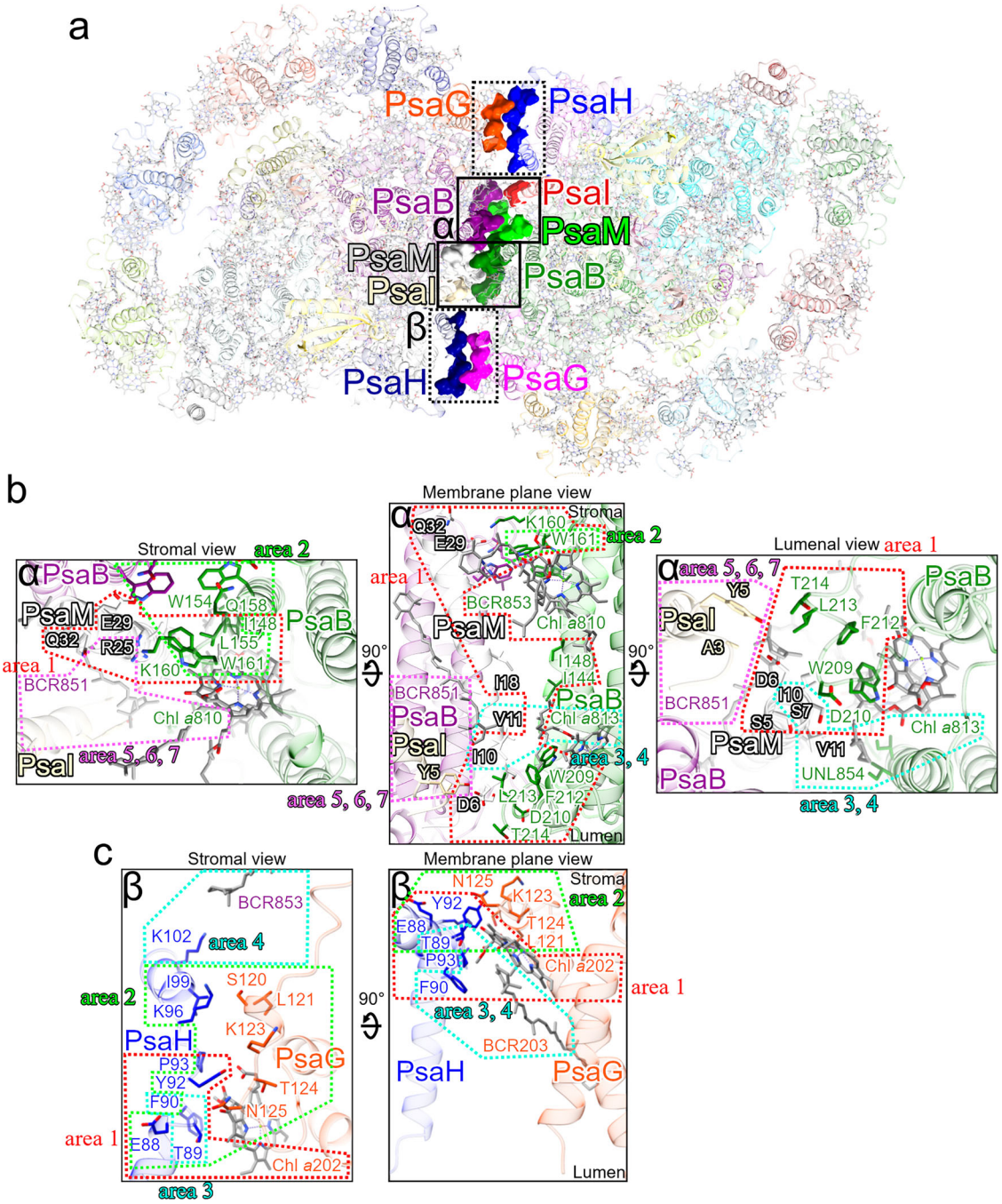


**Fig. 6 | The interface between the two monomers in a *Mp*PSI-LHCIs dimer. a** The monomer-monomer interface of the *Mp*PSI-LHCIs dimer indicated by squares. The surface of the amino acid residues participated in the interactions between the two monomers are listed in Supplementary Table 3 which is calculated by PISA analysis. The squares α with solid lines indicated the interface between PsaB, PsaI, and PsaM. The squares β with dotted lines indicated the interface between PsaG and PsaH. **b** The amino acid residues and ligands involved in the interactions between PsaB, PsaI, and PsaM viewed from the stromal side, lumenal side and along the membrane plane, respectively. **c** The amino acid residues and ligands involved in the interactions between PsaG and PsaH viewed from the stromal side and the along the membrane plane, respectively.

the one from *A. thaliana*, whereas *C. reinhardtii* has a longer BC loop than other species (boxed area with blue dotted lines in the bottom panel, blue square). Considering the conformation of Lhca4, it appears that *M. polymorpha* is closer to *A. thaliana*.

Unique features are found in the dimer of PSI from *M. polymorpha* compared with PSI dimers from other species The dimerization pattern of PSI in *M. polymorpha* is different from those in green algae, cyanobacteria and glaucophyte (Fig. 7). In *C. reinhardtii*, PSI dimerizes via Lhca9 with its associated co-factors[21] (Fig. 7b), whereas in cyanobacteria, PSI dimerizes via the stromal region of PsaB and PsaL[14–16], and trimerizes via the lumenal C-terminus of PsaL assisted by PsaM and PsaI[17,30,31] (Fig. 7c). In *C. paradoxa*, PSI tetramerizes via PsaA, PsaI, PsaK, PsaL, and PsaM[18] (Fig. 7d), which was obtained by adding a cross-linker into the sample. In our structure, the sample was isolated from the cells directly without a cross-linker, and there are no Lhcas existing at the PsaG/ PsaB/PsaM/PsaI/PsaH side, which form the dimerization interface in *Mp*PSI-LHCI. In *M. polymorpha*, PsaM and PsaH interact with PsaB and PsaG separately (Fig. 7a). At the interface near PsaB, PsaI and PsaM, the buried areas of both monomers are very similar and amount to 492.94 and 489.37 Å$^2$, respectively, and most of the buried areas are on the PsaB and PsaM subunits. PsaM is found in

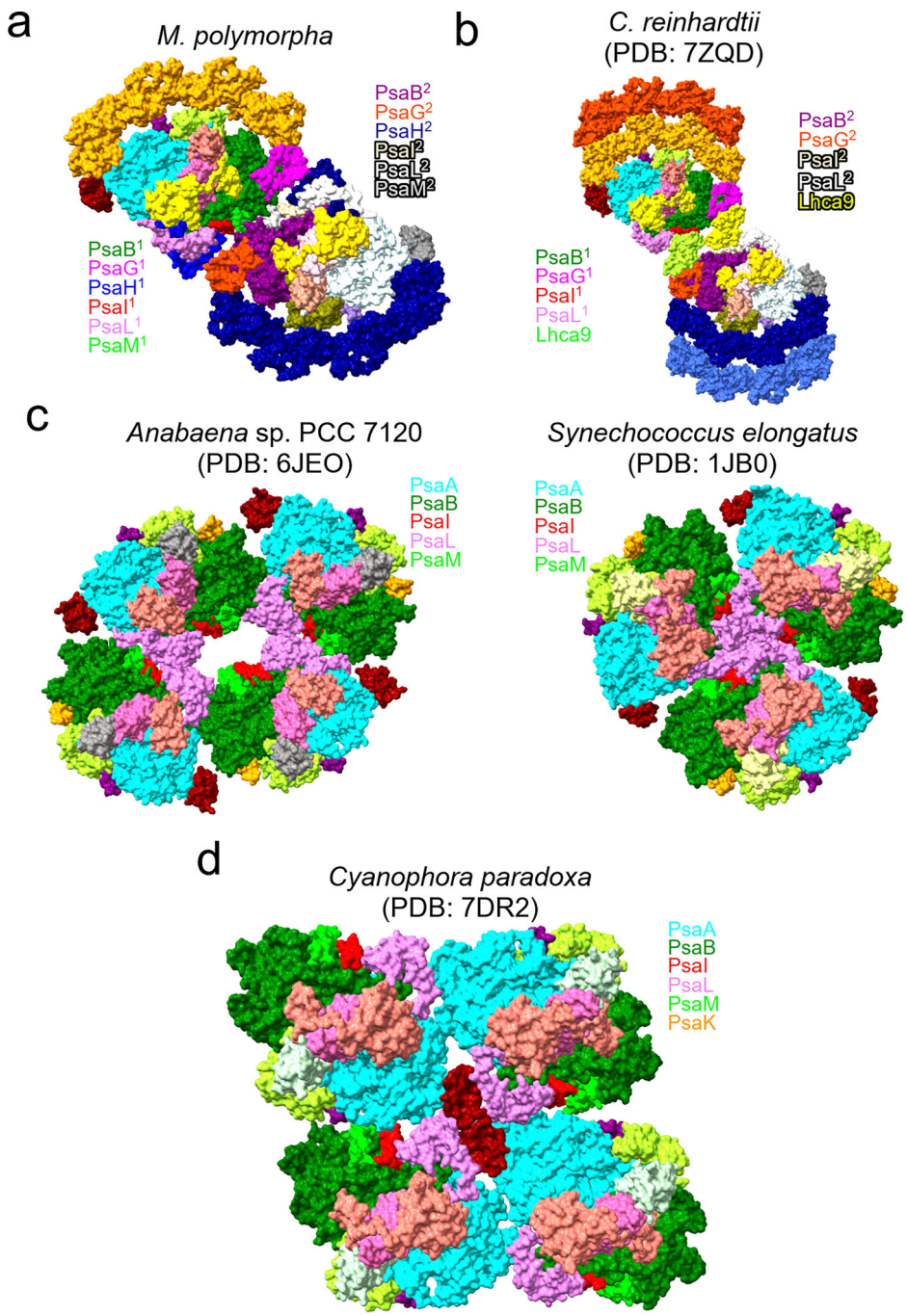


**Fig. 7 | A comformation for the oligomerization of photosystem I in different organisms. a** *M. polymorpha*. **b** *C. reinhardtii*. **c** Cyanobacteria *Anabaena* PCC 7120 and *S. elongatus*. **d** A glaucophyte *Cyanophora paradoxa*. PSI–LHCIs are oligomerized via PsaB, PsaG, PsaH, PsaI and PsaM in *Mp*PSI-LHCIs dimer, via Lhca9 in *Cr*PSI-LHCIs dimer, via PsaA, PsaB and PsaL in *Anabaena* PSI-LHCIs tetramer, and via PsaL and PsaI in *Se*PSI-LHCIs trimer, via PsaA, PsaI, PsaK, PsaL, and PsaM in *C. paradoxa* tetramer. The interactions between adjacent monomers are shown with black solid lines.

cyanobacteria[17], red algae[32–34], diatoms[35–37], cryptophytes[38,39], a haptophyte[40], and bryophytes[12,13,23], but not in the vascular plants[7–9]. PsaM is not found in the single-cell green algae *C. reinhardtii*[10,11] and *Dunalliela salina*[41], but is present in other single-cell green algae *Coccomyxa subellipsoidea*[25] and *Chlorella ohadii*[42], and the macroscopic green alga *Bryopsis corticulans*[43]. PsaM is involved in PSI trimerization in cyanobacteria[31], and in mediating the binding of an additional Lhca dimer in red algae and some green algae. In our *Mp*PSI-LHCI dimer structure, PsaM plays a key role in dimerization by interacting with PsaB and PsaM from the other monomer. These results illustrate the changes in the function of PsaM during evolution.

At the interface near PsaG and PsaH, the buried areas of the two monomers are 323.89 Å$^2$ and 358.11 Å$^2$, respectively, and most of the buried areas are at the stromal side of the PsaG and PsaH subunits. PsaG and PsaH are found in green algae, bryophytes and vascular plants, among which, PsaG is reported to be involved in the regulation of electron transport in *A. thaliana*[44,45], and PsaH is involved in PSI-LHCII trimer association during state transitions[46,47]. The density of the loop of the PsaH subunit is not very clear in the monomer structure, with 78 amino acid residues identified, but it is identified clearly in the dimer, with the structure of 90 amino acid residues solved. On the other hand, PsaL is not involved directly in the monomer-monomer interactions in the dimer in *M. polymorpha*, and the density of its loop is also not visible clearly in the monomer with 128 amino acid residues identified, but can be identified clearly in the dimer with 158 amino acid residues identified. These results indicate that PsaH and PsaL are more stable in the dimer, and PsaL may assist PsaH in the interactions with PsaG on the other monomer. There is no PsaM and PsaH in the previously reported structure of PSI-LHCI dimer from *C. reinhardtii*[21], which suggests that the dimerization mechanism is different among different species, and both PsaH and PsaM are essential for the dimerization of PSI-LHCI in bryophytes.

In summary, we determined the high-resolution cryo-EM structures of PSI-LHCI monomer and dimer from a liverwort *M. polymorpha*. In the electron transfer chain, the PsaB-Trp667 is rotated around 4°, and the distance between PsaB-Trp667 and $PhQ_A$ is around 0.5 Å longer compared to that from *P. patens*. Each PSI core binds 4 Lhcas, but there are differences in the positions of loops and Chls among the 4 Lhcas. Similar differences are observed in *P. patens* except for the shorter AC loop of Lhca4, which is observed only in *M. polymorpha*. The dimer structure represents the first PSI-LHCI dimer structure from land plants, in which PsaB, PsaM, PsaG, PsaI, and PsaH were found to play important roles in the dimerization. Compared with the dimer of *M. polymorpha*, cyanobacteria and green algae have different dimerization interfaces, and vascular plants may also have different interfaces because of the absence of PsaM in the PSI core complex.

## Methods

### Purification of PSI–LHCI from *M. polymorpha*

*M. polymorpha* gemmae of Takaragaike-1 (male) and Takaragaike-2 (female) accessions[48] were grown on a half-strength Gamborg's B5 medium supplemented with 0.5 g/L MES and 1% agar (pH was adjusted to 5.5 with KOH)[49,50] under continuous light at 10–15 μmol photons $m^{-2}s^{-1}$ at 20 °C. The thalli were harvested after 1 month of incubation, and thylakoid membranes were isolated as previously described with a few modifications[13,23]. The thalli were resuspended in a lysis buffer (30 mM Tricine-NaOH [pH 8.0], 15 mM NaCl, 10 μg/mL trypsin inhibitor, 10 μg/mL aprotinin, 10 μg/mL leupeptin) and homogenized with a blender (Warning; #37BL84[BL6]) for 30 s twice separated by 30 s. The homogenate was filtered through 16 layers of gauze, and the filtrate was centrifuged at 40,000 × *g* for 20 min. The pellet was resuspended in the lysis buffer as thylakoid membranes, and the following steps were carried out on ice.

To purify the *Mp*PSI-LHCI monomer supercomplex, the thylakoid membrane was diluted to a concentration of 0.6 mg Chl *a* $mL^{-1}$ and solubilized with 3.6% (w/v)α-DDM (Dojindo 347-06163) for 30 min on ice in the dark under gentle stirring. Insolubilized materials were removed by centrifugation at 40,000 × *g* for 20 min at 4 °C, and the supernatant was collected and loaded onto a 10–30% continuous sucrose gradient in the lysis buffer containing 0.03% α-DDM, which was centrifuged at maximum 230,000 × *g* for 16 h at 4 °C with a P40ST swing rotor (13 PA Tube 1.5 × 9.6 cm; Himac CP 80NX, HITACHI). The green band at around 20% concentration of sucrose was found to contain the *Mp*PSI-LHCI monomer supercomplex, which was collected (Supplementary Fig 2a). The sucrose was removed by dilution with a storage buffer (30 mM Tricine-NaOH [pH 8.0], 15 mM NaCl, 0.03% α-DDM), and the sample was concentrated with a 100 K cut-off concentrator (Amicon; UFC810024).

To isolate the *Mp*PSI-LHCI dimer supercomplex, the thylakoid membrane was diluted to a concentration of 0.6 mg Chl *a* $mL^{-1}$ and solubilized with 2.4% (w/v) α-DDM (Dojindo; 347-06163) for 30 min on ice in the dark under gentle stirring. Insolubilized materials were removed by centrifugation at 40,000 × *g* for 20 min at 4 °C, and the supernatant was collected and loaded onto a 15–30% continuous sucrose gradient in the lysis buffer containing 0.03% α-DDM. The green band at around 25% concentration of sucrose was collected (Supplementary Fig 2a), and the sucrose was removed in the same way as described in the *Mp*PSI-LHCI monomer purification.

### Biochemical analysis of the purified PSI–LHCI

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) containing 7.5 M urea was performed as previously described[51], with 20% acrylamide gel. The samples were denatured with 2% lithium lauryl sulfate, 60 mM dithiothreitol, and 60 mM Tris-HCl (pH 8.5) at 60 °C for 10 min before being loaded onto the SDS-PAGE. Clear native-PAGE (CN-PAGE) was performed as described previously[52–54]. The 3-color Prestained XL-Ladder (APRO; #SP-2140) and NativeMark™ Unstained Protein Standard (Invitrogen; #LC0725) were used as the protein molecular weight markers for SDS-PAGE and CN-PAGE, respectively. The absorbance spectra were measured with purified PSI-LHCI at 1 μg Chl *a* $mL^{-1}$ (for *Mp*PSI-LHCI monomer) and 7.5 μg Chl *a* $mL^{-1}$ (for *Mp*PSI-LHCI dimer) in the storage buffer from 400 to 750 nm at room temperature, which are normalized at the $Q_y$ peak.

### Mass spectrometry analysis of SDS-PAGE bands

For quantitative analysis by mass spectrometry, each band in the SDS-PAGE was cut out and digested by a trypsin digestion kit following the protocol provided by the company (Vanquish Neo, Thermo Scientific). Mass analysis was performed with an LC/QTOF mass spectrometer (LC: Nexera Mikros; MS: LCMS-9050, Shimadzu). Desalted samples were eluted with 50% acetonitrile containing 0.1% formic acid, and then injected into concentrated L-column2 ODS column (5.0 μm, 0.3 × 5 $mm^2$) after dilution to 5% acetonitrile containing 0.1% formic acid. Peptides were separated with a Shim-pack MC C18 column (1.9 μm, 0.175 × 50 mm; flowrate: 5 μL/min). The elution buffers used were 0.1% formic acid for buffer A and 100% acetonitrile containing 0.1% formic acid for buffer B. Peptides were eluted with a linear gradient from 5% to 45% buffer B for 20 min, including a regeneration step at 95% buffer B for 3 min, followed by an equilibration step at 5% buffer B. The samples were analyzed with QTOF mass in a positive ion modes. The proteins were identified using Mascot algorithm with the amino acid sequences of *M. polymorpha* and higher plants from the database.

### Pigment extraction and HPLC analysis

An isolated PSI-LHCI sample equivalent to 50 μg of Chl a was dissolved in 500 μL of 50:50 methanol/acetone (#21915-64, nacalai tesque; #014-08681, Fujifilm Wako) and centrifuged. The supernatant was dried under flow of nitrogen gas and resuspended in 500 μL of 60:40 acetone/ethyl acetate (#14746-91, nacalai tesque). After adding 400 μL of $H_2O$, the sample was centrifuged at 17,000 *g* for 5 min (120,000 rpm with T15A44 rotor; Himac CF 15RN, HITACHI) and the upper ethyl acetate layer was collected and dried under a flow of nitrogen gas, then resuspended in 200 μL 100% methanol. Before HPLC analysis, the resulting sample was filtered through a 0.45 μm PTFE filter, (Minisart RC 4, Millipore). Carotenoids and Chl pigments were analyzed using a Shimadzu HPLC system equipped with a Wakosil-II 5C18-100 column (5 μm, 250 ×4.6 $mm^2$, Wako) equipped with a Wakosil-II 5C18-100 guard column (5 μm, 10 ×4.6 $mm^2$, Wako), with 100% methanol as the eluent at a flow rate of 1.5 mL/min. Elution profiles and absorption spectra were monitored using an SPD-M20A photodiode array detector (Shimadzu, Japan), and the pigments were identified based of their absorption spectra (Supplementary Fig 12 and Supplementary Data 2).

### Negative-staining electron microscopy

For negative-staining EM, 4 μL of the sample diluted to 15 μg Chl *a* $mL^{-1}$ was applied to a glow-discharged carbon-coated copper grid (F-200 Cu, Nisshin EM), and the sample was incubated on the grid for 30 s. Excess sample solutions on the grid were removed using a filter paper, and the sample was stained with 4 μL of 8% ammonium molybdate for 30 s, and excess ammonium molybdate was removed with a filter paper. The grid was observed using a 200 kV JEM2100plus (JEOL Ltd.) electron microscope, recorded at 80,000 magnification.

### Cryo-electron microscopic data collection

For cryo-EM observation of the PSI-LHCI monomer, an aliquot of 4 μL of *Mp*PSI-LHCI monomer (1.5 mg Chl *a* $mL^{-1}$) was applied to a Quantifoil R1.2/1.3 Cu 300 mesh grid (Microtools GmbH). The grid was first hydrophilized by glow discharge using a JEC-3000FC auto fine coater (JEOL Ltd.) with 7 mA under 7 Pa pressure for 10 s. After sample loading, the grid was blotted once for 1 s with a force level of 1 under 100% humidity at 8 °C, and immediately plunge-frozen in liquid ethane by a Vitrobot Mark IV (Thermo Fisher Scientific). For cryo-EM observation of the PSI-LHCI dimer, an aliquot of 3 μL of *Mp*PSI-LHCI dimer (1.46 mg Chl *a* $mL^{-1}$) was applied to a grid treated with the same condition as above. After sample loading, the grid was blotted once for 6 s with a force level of 6 under 100% humidity at 8 °C,

and immediately plunge-frozen in liquid ethane by a Vitrobot Mark IV (Thermo Fisher Scientific).

The frozen grid was stored in liquid nitrogen before data collection. All micrographs were collected using an EPU2 software on a Krios G4 electron microscope (Thermo Fisher Scientific) equipped with a cold-field emission gun, an imaging filter (Selectris X, Thermo Fisher Scientific), and a direct electron detection camera (Falcon 4i, Thermo Fisher Scientific). The microscope was operated at 300 kV with a nominal magnification of 165 k.

In total, 9576 electron counting movies for *Mp*PSI-LHCI monomer and 8302 and 9605 electron counting movies for the first and second collection of *Mp*PSI-LHCI dimer were recorded in a pixel size of 0.727 Å, under a total electron dose of 50 electron $Å^{-2}$ for each movie. The two collections of *Mp*PSI-LHCI dimer were performed with 2 different grids blotted with the same condition as above. For *Mp*PSI-LHCI monomer, each exposure of 3.57 s was dose-fractionated into 1240 EER frames, and the nominal defocus range was set to −0.8 to −2.0 µm. For *Mp*PSI-LHCI dimer, each exposure of 3.69 and 3.98 s was dose-fractionated into 1240 EER frames, respectively, and the nominal defocus range was set to −0.6 to −1.8 µm for the first and second collection, respectively.

### Cryo-EM data processing

The workflow of cryo-EM data processing and model reconstruction was summarized in Supplementary Fig 3 and Supplementary Fig 4. Image processing was mainly performed using cryoSPARC 4.6.0[55]. The movie frames were aligned using the patch motion correction to obtain a final dose-weighted image[55]. Estimation of the contrast transfer function (CTF) was performed with patch-based CTF estimation[56]. For *Mp*PSI-LHCI monomer, a total of 1,078,038 particles were automatically picked from 9576 micrographs and were used for reference-free 2D classification and Ab-Initio analysis without imposing symmetry (Table 1). A total of 157,685 particles were selected from good classes of the ab-Initio analysis. Following this, 3D refinement and post-processing (global and local CTF refinement and reference motion correction) were performed, which yielded a 3D map at an overall resolution of 1.94 Å for the *Mp*PSI-LHCI monomer.

For *Mp*PSI-LHCI dimer, a total of 14,441 particles from 1 good class of ab-Initio analysis were selected from the first collection of 8302 micrographs, and were used to create template for template picking. The template was used for the first collection of 8302 micrographs and the second collection of 9605 micrographs, and a total of 968,224 and 1,196,413 particles were picked from the two collections, respectively. After multiple rounds of reference-free 2D classification and ab-Initio analysis without imposing symmetry, a total of 24,456 particles from the first collection and 14,747 particles from the second collection were selected. The particles from both collections were merged and classified by Ab-Initio analysis. Finally, a total of 27,750 particles were selected, and subsequent 3D refinement with C2 symmetry and post-processing (global and local CTF refinement and reference motion correction) yielded a 3D map at an overall resolution of 2.52 Å for the *Mp*PSI-LHCI dimer. The resolutions of the maps were estimated by the gold standard Fourier shell correlation (FSC) curve with a cut-off value of 0.143[57] (Supplementary Fig 4 and Supplementary Data 2).

### Model building and refinement

For model building, the structure of PSI-LHCI from *P. patens* (PDB ID: 6L35)[12] was manually fitted into the 1.94 Å *Mp*PSI-LHCI monomer map using UCSF ChimeraX v1.8[58]. The amino acid sequences of subunits were changed to those of *M. polymorpha*, and the model was manually built using Coot v0.9.8.7[59]. We searched the amino acid sequences of PSI core and Lhca from the sequences of *M. polymorpha*, and found 13 sequences of PSI core and 4 sequences of Lhca. The accession codes of subunits are listed in Supplementary Table 1. Chl *b* was assigned based on the density map as well as on the presence of hydrogen-bonding interactions between the C7-formyl group of Chl *b* and nearby residues, and those Chls with no clear interactions were modeled as Chl *a*. Automatic refinement was performed with phenix.real_space_refine of the Phenix v1.20.1 software suite[60]. Manual refinements and additions of water molecules were performed with Coot v0.9.8.7.

The final *Mp*PSI-LHCI monomer structure without water molecules was fitted into one side of the 2.52 Å *Mp*PSI-LHCI dimer cryo-EM map. Manual refinements were performed with Coot v0.9.8.7 as described above. Chl *b* in *Mp*PSI-LHCI dimer was assigned in the same way as for the *Mp*PSI-LHCI monomer. The model was duplicated and fitted into the other side of the map after all refinements were completed. The resolutions of the models were estimated by the gold standard model-map FSC curve with a cut-off value of 0.5[61] (Supplementary Fig 4 and Supplementary Data 2). Statistics of the structural analysis are summarized in Table 1.

### Interface analysis by PISA

The PISA interface analysis[27] was performed with the PSI dimer using only PsaB, PsaG, PsaH, PsaI, PsaJ, and PsaM subunits, including all the ligands (lipids and pigments) in the structure, with the automatic processing mode. The interfaces and interacting amino acids identified by this analysis are shown in Supplementary Table 3.

### FRET calculations

Förster resonance energy transfer (FRET) rate constants ($K_{FRET}$) were calculated based on the FRET theory from the formula $K_{FRET} = (Ck^2)/(n^4R^6)$[62–64], where $C$ is the factor calculated from the spectral overlap integral between two Chls, $k$ is the dipole orientation factor, $n$ is the refractive index and $R$ is the distance between the magnesium atoms of each Chls. The $C$ values applied for Chl *a* to Chl *a*, Chl *a* to Chl *b*, Chl *b* to Chl *a*, and Chl *b* to Chl *b* energy transfer were 32.26, 1.11, 9.61, and 14.45, respectively, and the $n$ value was 1.55, as described by Gradinaru et al.[63]. $k^2$ is defined as $k^2 = [\hat{u}_D \cdot \hat{u}_A - 3 \cdot (\hat{u}_D \cdot \hat{R}_{DA}) \cdot (\hat{u}_A \cdot \hat{R}_{DA})]^2$, where $\hat{u}_D$ and $\hat{u}_A$ are the transition dipole moment vectors of donor and acceptor Chls derived from the vectors of the coordinates of NB and ND atoms of Chls, respectively, and $\hat{R}_{DA}$ is the distance of the donor Chl magnesium to the acceptor Chl magnesium. FRET rates were computationally calculated using Kim's algorithm on the Python platform (Python v.3.10)[62]. The figures of the FRET pathway were drawn by UCSF ChimeraX v1.8[58].

### Phylogenetic analysis

The sequences of amino acid were downloaded from UniProt database (https://www.uniprot.org/). Alignment of amino acid sequences and construction of the phylogenetic tree were performed with CLC Main Workbench 25.0.3 (CLC Bio-Qiagen). The phylogenetic tree was generated using the unweighted pair group method using average linkages (UPMGA), Kimura Protein as the distance measure, and a bootstrap value of 1000 replicates.

### Statistics and reproducibility

The purification of PSI-LHCI monomer was performed at least 3 times with different lots of thalli, and PSI-LHCI dimer was purified twice with different lots of thalli. SDS-PAGE and CN-PAGE were performed for each purified PSI-LHCI sample. Mass spectrometry analysis and pigment analysis were performed once with PSI-LHCI monomer. The cryo-EM data for the monomer were collected from 1 grid, and the data for the dimer were collected from 2 grids on different dates. Data collection, refinement and validation statistics are summarized in Table 1.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Uncropped and unedited blot images were provided as Supplementary Fig. 13. All 9,576 raw cryo-EM movies and motion-corrected images for the *Mp*PSI-LHCI monomer supercomplex and all 17,907 raw cryo-EM movies for the *Mp*PSI-LHCI dimer supercomplex were deposited in the Electron Microscopy Public Image Archive (EMPIAR)[65] under the accession codes EMPIAR-12712 and EMPIAR-12713, respectively [https://www.ebi.ac.uk/empiar/EMPIAR-12712, https://www.ebi.ac.uk/empiar/EMPIAR-12713].

Atomic coordinates and cryo-EM maps were deposited in the Protein Data Bank under an accession code 9LUT and in the Electron Microscopy Data Bank[66] under an accession code EMD-63404 for the *Mp*PSI-LHCI monomer structure, and accession codes 9LUU and EMD-63405 for the *Mp*PSI-LHCI dimer structure.

Received: 11 July 2025; Accepted: 22 January 2026;
Published online: 05 February 2026

## References

1. Heathcote, P., Jones, M. R. & Fyfe, P. K. Type I photosynthetic reaction centres: structure and function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358**, 231–243 (2003).
2. Caspy, I. & Nelson, N. Structure of the plant photosystem I. *Biochem. Soc. Trans.* **46**, 285–294 (2018).
3. Qin, X. Structure, function, and evolution of photosystem I-light harvesting antenna I complexes. (eds Shen, J.-R., Satoh, K. & Allakhverdiev, S. I.) in *Photosynthesis: Molecular Approaches to Solar Energy Conversion*, 195–221 (Springer International Publishing, 2021).
4. Suga, M. & Shen, J.-R. Structural variations of photosystem I-antenna supercomplex in response to adaptations to different light environments. *Curr. Opin. Struct. Biol.* **63**, 10–17 (2020).
5. Pan, X., Cao, P., Su, X., Liu, Z. & Li, M. Structural analysis and comparison of light-harvesting complexes I and II. *Biochim. Biophys. Acta BBA Bioenerg.* **1861**, 148038 (2020).
6. Shen, J.-R. Structure, function, and variations of the photosystem I-antenna supercomplex from different photosynthetic organisms. in (eds Harris, J. R. & Marles-Wright, J.) *Macromolecular Protein Complexes IV: Structure and Function*, 351–377 (Springer International Publishing, 2022).
7. Ben-Shem, A., Frolow, F. & Nelson, N. Crystal structure of plant photosystem I. *Nature* **426**, 630–635 (2003).
8. Amunts, A., Drory, O. & Nelson, N. The structure of a plant photosystem I supercomplex at 3.4 Å resolution. *Nature* **447**, 58–63 (2007).
9. Qin, X., Suga, M., Kuang, T. & Shen, J.-R. Structural basis for energy transfer pathways in the plant PSI-LHCI supercomplex. *Science* **348**, 989–995 (2015).
10. Su, X. et al. Antenna arrangement and energy transfer pathways of a green algal photosystem-I–LHCI supercomplex. *Nat. Plants* **5**, 273–281 (2019).
11. Suga, M. et al. Structure of the green algal photosystem I supercomplex with a decameric light-harvesting complex I. *Nat. Plants* **5**, 626–636 (2019).
12. Yan, Q. et al. Antenna arrangement and energy-transfer pathways of PSI–LHCI from the moss Physcomitrella patens. *Cell Discov.* **7**, 10 (2021).
13. Sun, H., Shang, H., Pan, X. & Li, M. Structural insights into the assembly and energy transfer of the Lhcb9-dependent photosystem I from moss Physcomitrium patens. *Nat. Plants* **9**, 1347–1358 (2023).
14. Kato, K. et al. Structure of a cyanobacterial photosystem I tetramer revealed by cryo-electron microscopy. *Nat. Commun.* **10**, 4929 (2019).
15. Zheng, L. et al. Structural and functional insights into the tetrameric photosystem I from heterocyst-forming cyanobacteria. *Nat. Plants* **5**, 1087–1097 (2019).
16. Chen, M. et al. Distinct structural modulation of photosystem I and lipid environment stabilizes its tetrameric assembly. *Nat. Plants* **6**, 314–320 (2020).
17. Jordan, P. et al. Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* **411**, 909–917 (2001).
18. Kato, K. et al. Structure of a tetrameric photosystem I from a glaucophyte alga Cyanophora paradoxa. *Nat. Commun.* **13**, 1679 (2022).
19. Yadav, K. N. S. et al. Supercomplexes of plant photosystem I with cytochrome b6f, light-harvesting complex II and NDH. *Biochim. Biophys. Acta BBA Bioenerg.* **1858**, 12–20 (2017).
20. Wood, W. H. J. et al. Dynamic thylakoid stacking regulates the balance between linear and cyclic photosynthetic electron transfer. *Nat. Plants* **4**, 116–127 (2018).
21. Naschberger, A. et al. Algal photosystem I dimer and high-resolution model of PSI-plastocyanin complex. *Nat. Plants* **8**, 1191–1201 (2022).
22. Shimamura, M. Marchantia polymorpha: taxonomy, phylogeny and morphology of a model system. *Plant Cell Physiol.* **57**, 230–256 (2016).
23. Gorski, C. et al. The structure of the Physcomitrium patens photosystem I reveals a unique Lhca2 paralogue replacing Lhca4. *Nat. Plants* **8**, 307–316 (2022).
24. Keable, S. M. et al. Room temperature XFEL crystallography reveals asymmetry in the vicinity of the two phylloquinones in photosystem I. *Sci. Rep.* **11**, 21787 (2021).
25. Tsai, P.-C., Kato, K., Shen, J.-R. & Akita, F. Structural study of the chlorophyll between Lhca8 and PsaJ in an Antarctica green algal photosystem I-LHCI supercomplex revealed by its atomic structure. *Biochim. Biophys. Acta BBA Bioenerg.* **1866**, 149543 (2025).
26. Punjani, A. & Fleet, D. J. 3D variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
27. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
28. Amunts, A., Ben-Shem, A. & Nelson, N. Solving the structure of plant photosystem I—biochemistry is vital. *Photochem. Photobiol. Sci.* **4**, 1011–1015 (2005).
29. Amunts, A., Toporik, H., Borovikova, A. & Nelson, N. Structure determination and improved model of plant photosystem I*. *J. Biol. Chem.* **285**, 3478–3486 (2010).
30. Chitnis, V. P. & Chitnis, P. R. PsaL subunit is required for the formation of photosystem I trimers in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett.* **336**, 330–334 (1993).
31. Naithani, S., Hou, J.-M. & Chitnis, P. R. Targeted inactivation of the psaK1, psaK2 and psaM genes encoding subunits of Photosystem I in the cyanobacterium Synechocystis sp. PCC 6803. *Photosynth. Res.* **63**, 225–236 (2000).
32. Pi, X. et al. Unique organization of photosystem I–light-harvesting supercomplex revealed by cryo-EM from a red alga. *Proc. Natl. Acad. Sci. USA* **115**, 4423–4428 (2018).
33. Antoshvili, M., Caspy, I., Hippler, M. & Nelson, N. Structure and function of photosystem I in Cyanidioschyzon merolae. *Photosynth. Res.* **139**, 499–508 (2019).
34. Kato, K. et al. The structure of PSI-LHCI from *Cyanidium caldarium* provides evolutionary insights into conservation and diversity of red-lineage LHCs. *Proc. Natl. Acad. Sci. USA* **121**, e2319658121 (2024).
35. Nagao, R. et al. Structural basis for assembly and function of a diatom photosystem I-light-harvesting supercomplex. *Nat. Commun.* **11**, 2481 (2020).
36. Xu, C. et al. Structural basis for energy transfer in a huge diatom PSI-FCPI supercomplex. *Nat. Commun.* **11**, 5081 (2020).
37. Kato, K. et al. Structural basis for molecular assembly of fucoxanthin chlorophyll a/c-binding proteins in a diatom photosystem I supercomplex. *eLife* **13**, RP99858 (2024).
38. Zhao, L.-S. et al. Structural basis and evolution of the photosystem I–light-harvesting supercomplex of cryptophyte algae. *Plant Cell* **35**, 2449–2463 (2023).
39. Zhang, S. et al. Growth phase-dependent reorganization of cryptophyte photosystem I antennae. *Commun. Biol.* **7**, 560 (2024).
40. He, F.-Y. et al. Structural insights into the assembly and energy transfer of haptophyte photosystem I–light-harvesting supercomplex. *Proc. Natl. Acad. Sci. USA* **121**, e2413678121 (2024).

41. Caspy, I. et al. Structure and energy transfer pathways of the Dunaliella Salina photosystem I supercomplex. *Biochim. Biophys. Acta BBA Bioenerg.* **1861**, 148253 (2020).
42. Caspy, I. et al. Cryo-EM photosystem I structure reveals adaptation mechanisms to extreme high light in Chlorella ohadii. *Nat. Plants* **7**, 1314–1322 (2021).
43. Qin, X. et al. Structure of a green algal photosystem I in complex with a large number of light-harvesting complex I subunits. *Nat. Plants* **5**, 263–272 (2019).
44. Jensen, P. E., Rosgaard, L., Knoetzel, J. & Vibe Scheller, H. Photosystem I activity is increased in the absence of the PSI-G Subunit*. *J. Biol. Chem.* **277**, 2798–2803 (2002).
45. Varotto, C. et al. Single and double knockouts of the genes for photosystem I subunits G, K, and H of arabidopsis. Effects on photosystem I composition, photosynthetic electron flow, and state transitions. *Plant Physiol.* **129**, 616–624 (2002).
46. Lunde, C., Jensen, P. E., Haldrup, A., Knoetzel, J. & Scheller, H. V. The PSI-H subunit of photosystem I is essential for state transitions in plant photosynthesis. *Nature* **408**, 613–615 (2000).
47. Jensen, P. E. et al. Structure, function and regulation of plant photosystem I. *Biochim. Biophys. Acta BBA Bioenerg.* **1767**, 335–352 (2007).
48. Ishizaki, K., Chiyoda, S., Yamato, K. T. & Kohchi, T. Agrobacterium-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. *Plant Cell Physiol.* **49**, 1084–1091 (2008).
49. Gamborg, O. L., Miller, R. A. & Ojima, K. Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell Res.* **50**, 151–158 (1968).
50. Ishizaki, K., Nonomura, M., Kato, H., Yamato, K. T. & Kohchi, T. Visualization of auxin-mediated transcriptional activation using a common auxin-responsive reporter system in the liverwort Marchantia polymorpha. *J. Plant Res.* **125**, 643–651 (2012).
51. Ikeuchi, M. & Inoue, Y. A New 4.8-kDa polypeptide intrinsic to the PS II reaction center, as revealed by modified SDS-PAGE with improved resolution of low-molecular-weight proteins. *Plant Cell Physiol.* https://doi.org/10.1093/oxfordjournals.pcp.a077628 (1988).
52. Schägger, H., Cramer, W. A. & von Jagow, G. Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* https://doi.org/10.1006/abio.1994.1112 (1994).
53. Kawakami, K., Iwai, M., Ikeuchi, M., Kamiya, N. & Shen, J.-R. Location of PsbY in oxygen-evolving photosystem II revealed by mutagenesis and X-ray crystallography. *FEBS Lett.* **581**, 4983–4987 (2007).
54. Wittig, I. & Schägger, H. Features and applications of blue-native and clear-native electrophoresis. *PROTEOMICS* **8**, 3974–3990 (2008).
55. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
56. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol*. **192**, 216–221 (2019).
57. Grigorieff, N. & Harrison, S. C. Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Curr. Opin. Struct. Biol.* **21**, 265–273 (2011).
58. Pettersen, E. F. et al. UCSF CHIMERAX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
59. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
60. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877 (2019).
61. Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 814–840 (2018).
62. Mazor, Y., Borovikova, A., Caspy, I. & Nelson, N. Structure of the plant photosystem I supercomplex at 2.6 Å resolution. *Nat. Plants* **3**, 17014 (2017).
63. Gradinaru, C. C. et al. The flow of excitation energy in LHCII monomers: implications for the structural model of the major plant antenna. *Biophys. J.* **75**, 3064–3077 (1998).
64. Eunchul K. Computational analysis of Förster resonance energy transfer in photosynthetic proteins (v1.0.0). Zenodo https://zenodo.org/records/3250649 (2019).
65. Iudin, A. et al. EMPIAR: the electron microscopy public image archive. *Nucleic Acids Res.* **51**, D1503–D1511 (2023).
66. The wwPDB Consortium et al EMDB—the Electron Microscopy Data Bank. *Nucleic Acids Res.* **52**, D456–D465 (2024).


## Acknowledgements

We thank Professor Takayuki Kohchi (Kyoto Univ.) for providing Tak-1 and Tak-2 accessions. This research was supported by JSPS KAKENHI No. JP22H04916 (J.-R.S.), JP25K08923 (F.A.) and JP24K02025 (H.M.), and a JSPS Program for Forming Japan's Peak Research Universities (J-PEAKS) Grant Number JPJS00420230010. This work was also supported by the Core-Facility Portal (CFPOU) at Okayama University (RIIS-n01 and CFPOU DGP_12), and we thank the International Research Center for Structural Biology of Okayama University for helping collection of the cryo-EM data.

## Author contributions

Pi-Cheng Tsai: Conceptualization, methodology, validation, investigation, visualization, writing—original draft. Romain La Rocca: Methodology, investigation, visualization. Hiroyasu Motose: Resources, funding acquisition. Jian-Ren Shen: Resources, writing—review & editing, funding acquisition, project administration, supervision. Fusamichi Akita: Conceptualization, resources, writing—review & editing, funding acquisition, project administration, visualization, supervision.

## Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s42003-026-09631-w.

**Correspondence** and requests for materials should be addressed to Jian-Ren Shen or Fusamichi Akita.

**Peer review information** *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary handling editors: Xiaoling Xu and Laura Rodríguez Pérez. A peer review file is available.