



# Structural insights into cyanobacterial photosystem II intermediates associated with Psb28 and Tsl0063

Yanan Xiao<sup>1,2,6</sup>, Guoqiang Huang<sup>3,6</sup>, Xin You<sup>3</sup>, Qingjun Zhu<sup>1,2</sup>, Wenda Wang<sup>1</sup>, Tingyun Kuang<sup>1</sup>, Guangye Han<sup>1</sup>✉, Sen-Fang Sui<sup>3,4</sup>✉ and Jian-Ren Shen<sup>1,5</sup>✉

**Photosystem II (PSII) is a multisubunit pigment-protein complex and catalyses light-induced water oxidation, leading to the conversion of light energy into chemical energy and the release of dioxygen. We analysed the structures of two Psb28-bound PSII intermediates, Psb28-RC47 and Psb28-PSII, purified from a *psbV*-deletion strain of the thermophilic cyanobacterium *Thermosynechococcus vulcanus*, using cryo-electron microscopy. Both Psb28-RC47 and Psb28-PSII bind one Psb28, one Tsl0063 and an unknown subunit. Psb28 is located at the cytoplasmic surface of PSII and interacts with D1, D2 and CP47, whereas Tsl0063 is a transmembrane subunit and binds at the side of CP47/PsbH. Substantial structural perturbations are observed at the acceptor side, which result in conformational changes of the quinone ( $Q_B$ ) and non-haem iron binding sites and thus may protect PSII from photodamage during assembly. These results provide a solid structural basis for understanding the assembly process of native PSII.**

Photosystem II (PSII) is a huge membrane protein complex located in the thylakoid membranes of oxygenic photosynthetic organisms, and it performs light-induced electron transfer reactions leading to the splitting of water into protons and molecular oxygen<sup>1,2</sup>. A functional PSII core from cyanobacteria consists of 20 protein subunits, including 17 intrinsic, transmembrane protein subunits and three membrane-extrinsic subunits as well as around 80 cofactors<sup>2–4</sup>. The structures of the cyanobacterial and red algal PSII dimers have been solved by X-ray crystallography at atomic resolutions<sup>3–5</sup>, and the structures of the PSII core in complex with its light-harvesting complexes from various organisms have recently been solved by cryo-electron microscopy (cryo-EM)<sup>6–10</sup>. However, the dynamic structure of PSII, such as the assembly and repair intermediates of this sophisticated protein machinery, remain largely unknown.

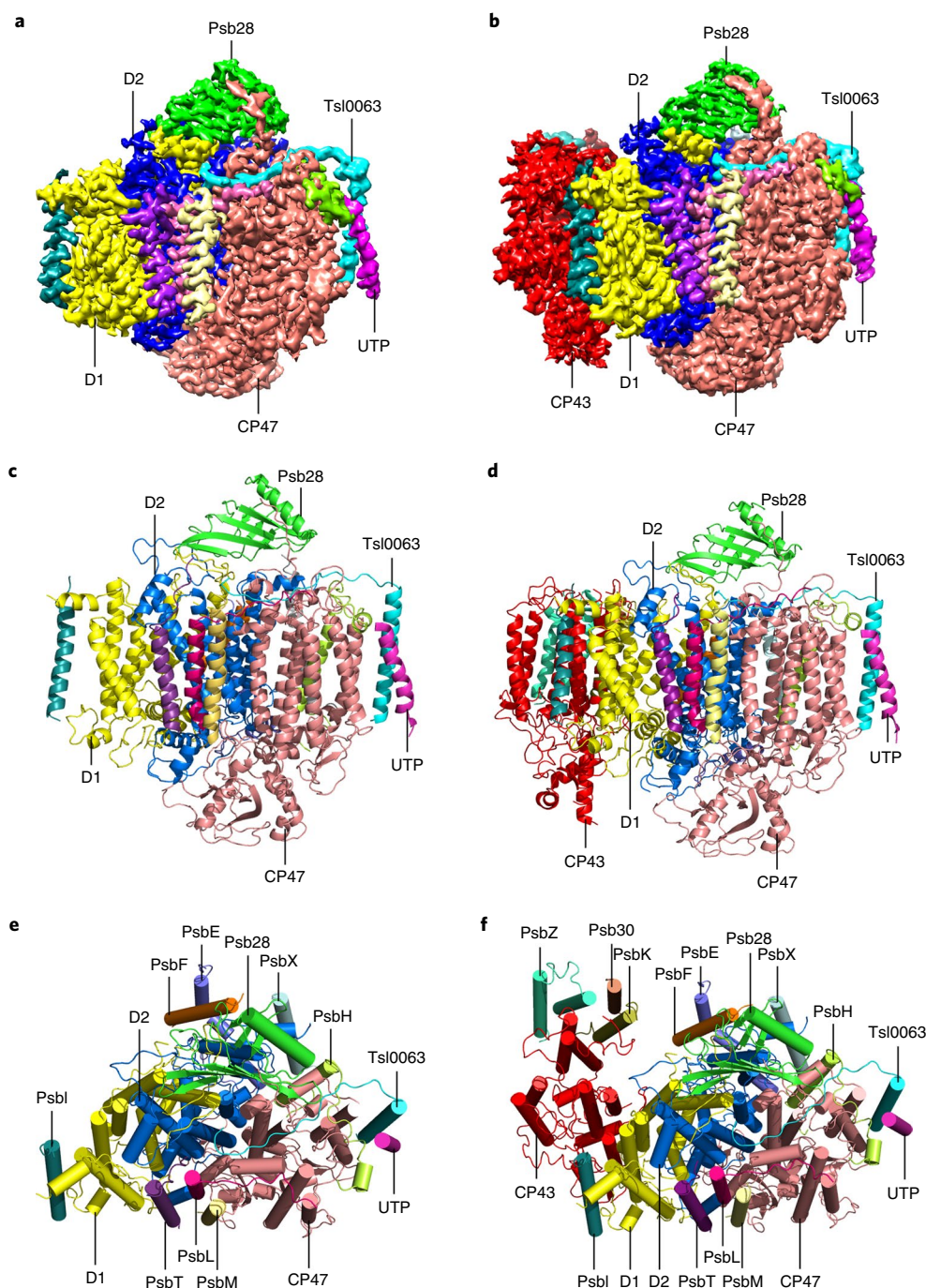
The assembly of PSII is a stepwise process involving several assembly modules and intermediate complexes. It begins with the formation of the PSII–reaction centre (RC) complex from the D1 and the D2 modules, followed by the attachment of the CP47 module to form the intermediate RC47, a subcomplex lacking the CP43 module<sup>11–14</sup>. The addition of the CP43 module to RC47 completes the assembly of transmembrane subunits and leads to the formation of apo-PSII, an intermediate that does not exhibit the water oxidation activity<sup>11–14</sup>. Several different apo-PSII structures have been revealed by recent studies, showing the structural dynamics of the protein assembly relevant to the  $Mn_4CaO_5$  cluster as well as other subunits<sup>15–19</sup>.

The assembly of the water-oxidizing catalyst  $Mn_4CaO_5$  cluster occurs through a light-driven process called photoactivation<sup>20</sup>. This process involves stepwise oxidation and incorporation of the Mn

ions as well as the binding of the Ca ion into the apo-PSII to form a full  $Mn_4CaO_5$  cluster. Three extrinsic proteins (PsbO, PsbU and PsbV in cyanobacteria) then attach to the lumenal surface and act as a cap to protect and stabilize the  $Mn_4CaO_5$  cluster, resulting in the final, active PSII core. During the assembly and/or repair of PSII, there are a number of auxiliary subunits that associate with the intermediate PSII complex and are detached from the final functional PSII once it is matured<sup>11,14,21–23</sup>. One such component is Psb27, a small lumen-localized protein associated with the inactive PSII<sup>24,25</sup>. This protein has been shown to facilitate the assembly of the  $Mn_4CaO_5$  cluster by preventing the binding of extrinsic proteins (PsbO, PsbU and PsbV) and stabilizing a structural arrangement for photoactivation<sup>18,19,25,26</sup>.

Another similar auxiliary subunit is Psb28, which is a soluble protein with a molecular weight of 13 kDa and is found in cyanobacteria and chloroplasts<sup>11,14,19,21,23,27</sup>. In *Synechocystis* sp. PCC 6803, there are two Psb28 homologues, Psb28-1 (encoded by *slr1398*) and Psb28-2 (encoded by *slr1739*)<sup>13,24</sup>. Psb28-1 (also called Psb13 or Ycf79) was originally identified as a peripheral protein associated with PSII in a substoichiometric amount<sup>24</sup>; it was subsequently shown to be an assembly cofactor involved in the biogenesis of PSII and maintenance of PSII activity during heat stress<sup>28,29</sup>. Psb28-1, instead of Psb28-2, preferentially binds to the cytoplasmic surface of PSII close to the CP47 and PsbH subunits, although both Psb28-1 and Psb28-2 associate with the monomeric PSII and RC47 complex<sup>13,14,27–30</sup>. Homologues of Psb28 were also found in algae and vascular plants, but few functional studies have been conducted<sup>27</sup>, except that the knockout of Psb28 (encoded by the gene *Os01g71190*) in *Oryza sativa* leads to a pale green phenotype<sup>31</sup>.

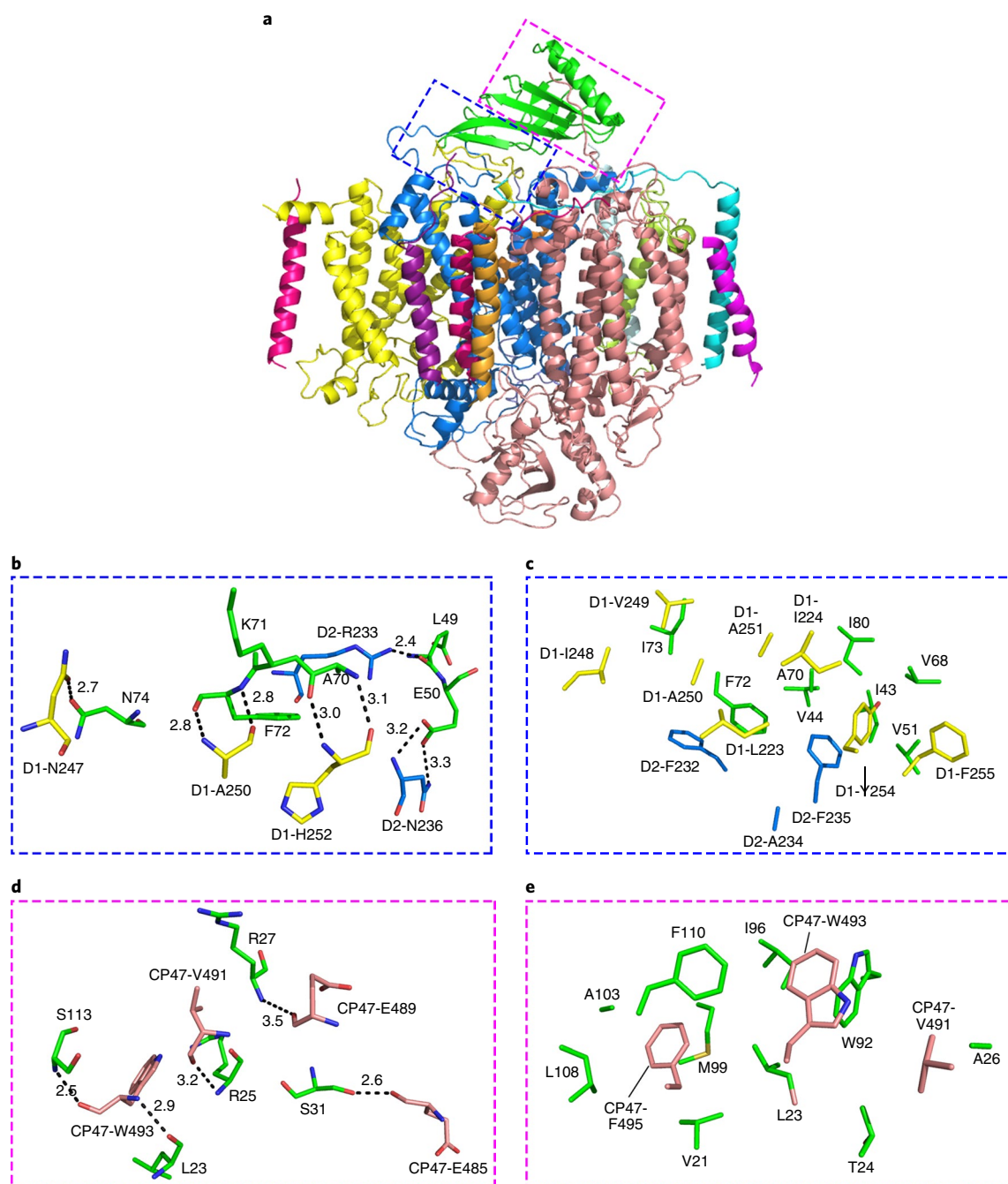
<sup>1</sup>Photosynthesis Research Center, Key Laboratory of Photobiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China. <sup>2</sup>University of Chinese Academy of Sciences, Beijing, China. <sup>3</sup>State Key Laboratory of Membrane Biology, Beijing Advanced Innovation Center for Structural Biology & Frontier Research Center for Biological Structure, School of Life Sciences, Tsinghua University, Beijing, China. <sup>4</sup>Department of Biology, Southern University of Science and Technology, Shenzhen, China. <sup>5</sup>Research Institute for Interdisciplinary Science and Graduate School of Natural Science and Technology, Okayama University, Okayama, Japan. <sup>6</sup>These authors contributed equally to this work: Yanan Xiao, Guoqiang Huang. ✉e-mail: [hanguangye@ibcas.ac.cn](mailto:hanguangye@ibcas.ac.cn); [suisf@mail.tsinghua.edu.cn](mailto:suisf@mail.tsinghua.edu.cn); [shen@cc.okayama-u.ac.jp](mailto:shen@cc.okayama-u.ac.jp)



**Fig. 1 | Overall structures of the Psb28-RC47 and Psb28-PSII complexes from *T. vulcanus*.** **a,b**, Surface representations of the Psb28-RC47 (**a**) and Psb28-PSII (**b**) monomers viewed along the membrane plane. **c,d**, Side views of the Psb28-RC47 (**c**) and Psb28-PSII (**d**) monomers along the membrane plane. **e,f**, Top views of the Psb28-RC47 (**e**) and Psb28-PSII (**f**) monomers from the stromal side.

In the past few years, many different types of PSII intermediates associated with Psb28 have been purified from cyanobacteria in a number of different biochemical preparations unable to evolve oxygen<sup>13,28–30,32–36</sup>. Intriguingly, in some of these preparations<sup>13,32</sup>, Psb28 and CP47 were co-purified with stress-induced small chlorophyll *a/b*-binding (CAB)-like proteins (SCPs), which are also called high-light-induced proteins (HLIPs) and are associated most prominently with the monomeric form of PSII<sup>37–39</sup>. These findings suggest that Psb28 might stabilize the association of SCPs with CP47 in monomeric PSII complexes under stress conditions.

The structure of isolated Psb28 has been solved by both nuclear magnetic resonance (NMR)<sup>40</sup> and X-ray crystallography<sup>41</sup>, and its association with PSII assembly intermediates has been studied by various techniques<sup>28,35</sup>. These studies have suggested that Psb28 binds to the cytoplasmic surface of PSII near the  $\alpha$ - and  $\beta$ -subunits of cytochrome *b*<sub>559</sub> and PsbH. However, the exact binding site and functions of Psb28 are unknown. Very recently, the structure of a PSII assembly intermediate containing Psb27, Psb28 and Tsl0063 from a *psbJ*-deletion mutant was solved by cryo-EM, showing the binding sites of Psb28 and Tsl0063 (a small transmembrane polypeptide that is sometimes termed Psb34)<sup>19</sup>.



**Fig. 2 | Interactions between Psb28 and the PSII core in Psb28-RC47.** **a**, Side view of the whole Psb28-RC47 complex, showing the binding site of Psb28 (highlighted with dotted boxes). **b,c**, Enlarged views of the binding area (blue box in **a**), showing the presence of possible hydrogen bonds (**b**) and hydrophobic interactions (**c**) between Psb28 and the D1 and D2 subunits. **d,e**, Enlarged views of the of binding area (magenta box in **a**), showing the presence of possible hydrogen bonds (**d**) and hydrophobic interactions (**e**) between Psb28 and CP47. The amino acid residues involved in the interactions are shown as sticks and coloured green in Psb28, yellow in D1, marine in D2 and salmon in CP47. The possible hydrogen bonds in **b** and **d** are shown in dashed lines, and the numbers near the dashed lines indicate bond lengths in Å.

To examine the binding site and functions of Psb28 in the assembly and repair of PSII more extensively, we purified Psb28-associated RC47 and PSII assembly intermediates from a PsbV-deletion mutant strain and a wild-type (WT) strain carrying His-tagged Psb28 from a thermophilic cyanobacterium, *Thermosynechococcus vulcanus*, and analysed their structures by single-particle cryo-EM at a resolution of 3.14 Å for both intermediates. The location and binding features of Psb28 and Tsl0063, together with an unidentified subunit, within these two PSII intermediates are revealed, which shed light

on the regulatory roles of Psb28 and Tsl0063 in the PSII assembly and repair cycle.

## Results

### Purification and properties of Psb28-RC47 and Psb28-PSII.

The Psb28-containing PSII complex was isolated by a Ni-affinity column from a HisPsb28-ΔPsbV strain lacking the extrinsic PsbV subunit (Supplementary Fig. 1 and Supplementary Table 1), since this strain gave rise to a higher amount of PSII assembly/repair

intermediates having Psb28 than that of the WT<sup>18,42</sup>. Three peaks were eluted from a Ni-affinity column of *n*-dodecyl- $\beta$ -D-maltoside ( $\beta$ -DDM)-solubilized thylakoid membranes (Extended Data Fig. 1a). Analyses by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), blue-native PAGE (BN-PAGE) and western blotting showed that fraction III was a Psb28-enriched PSII preparation, which consists of D1, D2, CP47, CP43, His-Psb28 and some small PSII subunits (Extended Data Fig. 1b–d). Notably, an additional PSII assembly cofactor, the product of a hypothetical open reading frame *tsl0063*, was detected in fraction III (Extended Data Fig. 1d), but three extrinsic subunits (PsbO, PsbV and PsbU) are absent in this fraction. BN-PAGE showed the presence of two major bands, band-1 and band-2, in this fraction. Band-1 migrates to the same position as that of PSII monomer, whereas band-2 has a molecular weight lower than that of band-1, suggesting that fraction III is a mixture of two monomeric PSII intermediates (Extended Data Fig. 1c). Among these bands, band-1 is composed of D1, D2, CP47, CP43, Psb28 and some small PSII subunits, whereas band-2 lacks the CP43 subunit, as confirmed by two-dimensional (2D) electrophoresis (Extended Data Fig. 1g). We therefore named the complexes in band-1 and band-2 as Psb28–PSII and Psb28–RC47, respectively.

The mixture of Psb28–PSII and Psb28–RC47 complexes in fraction III was further separated by an anion exchange column (Extended Data Fig. 2a), from which four fractions were obtained (fractions 1–4). Analyses by SDS-PAGE and BN-PAGE showed that fractions 1–3 are enriched in Psb28–PSII, whereas fraction 4 is enriched in Psb28–RC47 (Extended Data Fig. 2b,c). However, the two types of complexes were not completely separated; Psb28–PSII was predominantly found in the earlier fractions, whereas Psb28–RC47 was more abundant in the later fractions.

The Psb28-associated PSII complex was also isolated from the HisPsb28-WT strain and characterized (Extended Data Fig. 1e–g). The elution pattern of the HisPsb28-WT strain from the Ni-affinity column is similar to that of the HisPsb28- $\Delta$ PsbV strain, exhibiting three peaks (Extended Data Fig. 1e). Among them, fraction III contained Psb28-associated PSII and RC47, which had similar compositions of protein subunits as that found from HisPsb28- $\Delta$ PsbV (Extended Data Fig. 1f,g). This reveals that there are numerous PSII subparticles involved in the assembly process in both  $\Delta$ PsbV and WT strains. However, the relative contents of Psb28–PSII and Psb28–RC47 were different in fraction III. More Psb28–PSII was purified from HisPsb28- $\Delta$ PsbV, whereas Psb28–RC47 was more abundant than Psb28–PSII in the HisPsb28-WT strain (Extended Data Fig. 1c,f). Additionally, Psb27, an assembly/repair cofactor, was observed in a substoichiometric amount in fraction III of the HisPsb28-WT strain but not in the HisPsb28- $\Delta$ PsbV strain (Extended Data Fig. 1g).

#### Overall structure of the Psb28–RC47 and Psb28–PSII complexes.

We collected 3,461 and 2,207 cryo-EM micrographs of the monomeric, Psb28-associated PSII complexes Psb28–RC47 and Psb28–PSII purified from the HisPsb28- $\Delta$ PsbV strain (Supplementary Figs. 2 and 3). After several rounds of two-dimensional (2D) and 3D classification (Supplementary Figs. 2 and 3), 241,790 and 194,738 particles were selected for final 3D auto-refine, which yielded a cryo-EM map with an overall resolution of 3.14 Å for both the Psb28–RC47 and Psb28–PSII complexes (Supplementary Table 2).

Both Psb28–RC47 and Psb28–PSII are monomeric intermediates containing 14 and 18 protein subunits, respectively (Fig. 1 and Extended Data Fig. 3). In the Psb28–RC47 complex (Fig. 1a,c,e), 11 of these 14 subunits include 3 large intrinsic transmembrane subunits (D1, D2 and CP47) and 8 low-molecular-mass intrinsic transmembrane subunits (PsbE, PsbF, PsbH, PsbI, PsbL, PsbM, PsbT and PsbX), which are present in the native PSII<sup>3,4</sup>. On the basis of the cryo-EM density, we were able to build the structure of the extrinsic subunit Psb28, which is located at the stromal

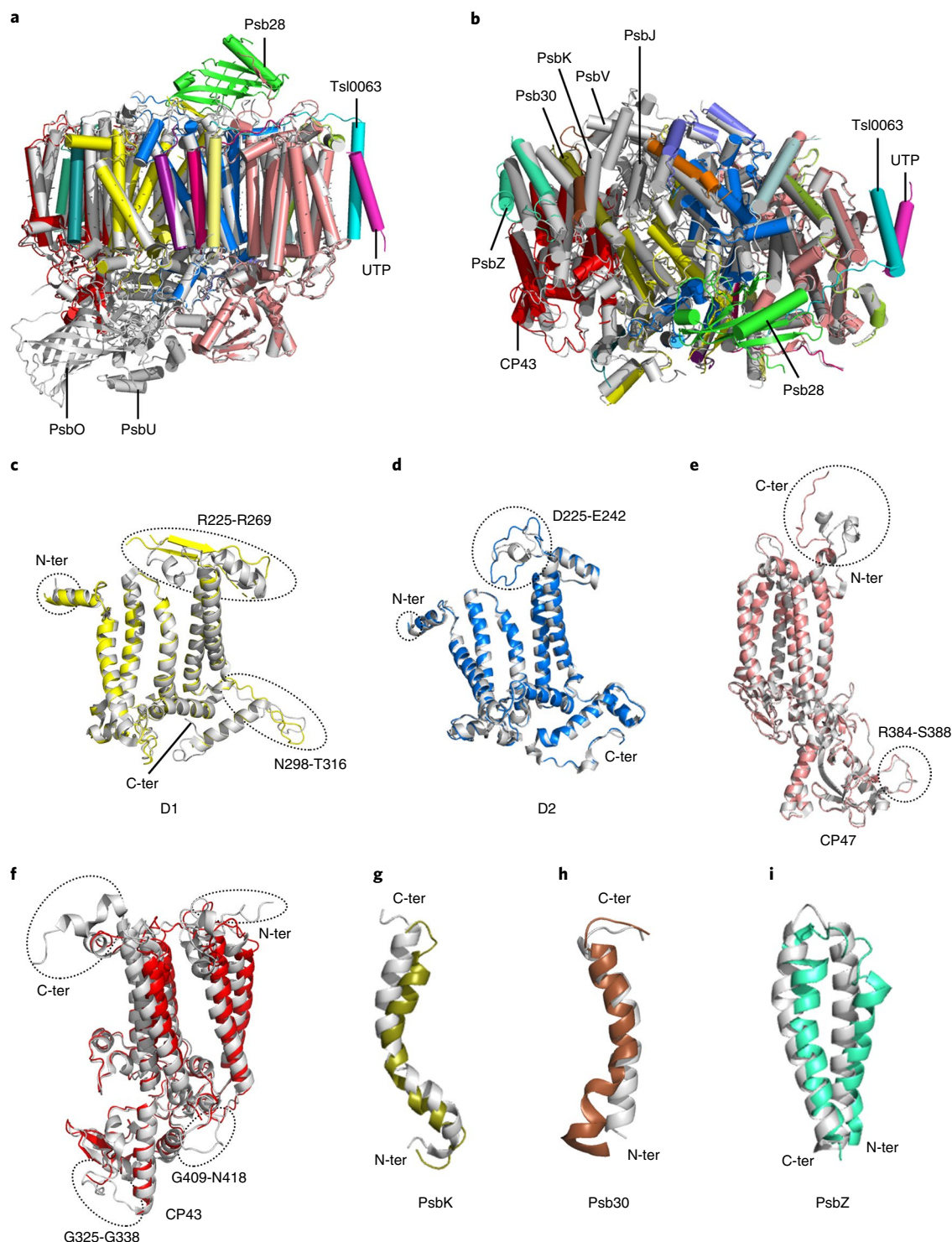
**Table 1 | Hydrogen bonds involved in the interactions between the Psb28 and Tsl0063 subunits and the PSII core in Psb28–RC47**

| Amino acid residue 1 | Amino acid residue 2 | Distance (Å) |
|----------------------|----------------------|--------------|
| Psb28-Ala70          | D1-His252            | 3.0, 3.1     |
| Psb28-Phe72          | D1-Ala250            | 2.8          |
| Psb28-Asn74          | D1-Asn247            | 2.7          |
| Psb28-Leu49          | D2-Arg233            | 2.4          |
| Psb28-Glu50          | D2-Asn236            | 3.2, 3.3     |
| Psb28-Leu23          | CP47-Trp493          | 2.9          |
| Psb28-Arg25          | CP47-Val491          | 3.2          |
| Psb28-Arg27          | CP47-Glu489          | 3.5          |
| Psb28-Ser31          | CP47-Glu485          | 2.6          |
| Psb28-Ser113         | CP47-Trp493          | 2.5          |
| Tsl0063-Glu7         | D1-Ile224            | 3.4          |
| Tsl0063-Leu11        | CP47-Gly2            | 3.1          |
| Tsl0063-Leu11        | CP47-Arg7            | 2.9          |
| Tsl0063-Asn13        | PsbL-Gln8            | 3.2, 3.3     |
| Tsl0063-Asn13        | CP47-Arg472          | 3.2          |
| Tsl0063-Ala15        | CP47-Asn14           | 3.5          |
| Tsl0063-Ala15        | CP47-Arg476          | 3.1          |
| Tsl0063-Glu17        | CP47-Arg476          | 3.3          |
| Tsl0063-Tyr21        | PsbH-Gly19           | 3.2          |
| Tsl0063-Tyr21        | PsbH-Val21           | 2.9          |
| Tsl0063-Ala23        | PsbH-Val21           | 2.7          |

surface over the top of D1, D2 and CP47. We also identified densities corresponding to two protein subunits with one transmembrane helix each, marked as Tsl0063 and unidentified transmembrane protein (UTP). The Tsl0063 subunit is encoded by the gene *tsl0063* and has sequence similarity to proteins of the CAB/ELIP/HLIP and light-harvesting-protein-like family<sup>39</sup> (Extended Data Fig. 4). However, no chlorophylls are found in Tsl0063 in the Psb28–RC47 and Psb28–PSII complexes. The second transmembrane helix protein subunit, denoted UTP, has a lower quality of cryo-EM density, so its sequences were not identified and were tentatively assigned as poly-alanines in the structure. It may be a second copy of Tsl0063 but may also be a different subunit.

The Psb28–PSII complex contains all subunits that are found in Psb28–RC47 with the addition of the CP43, PsbK, Psb30 and PsbZ subunits (Fig. 1b,d,f). The root mean square deviation values of the  $\alpha$  backbone between the Psb28–RC47 and Psb28–PSII cores (not including the four subunits present in Psb28–PSII only), Psb28–RC47 and the monomer of native PSII (PDB: 3WU2) (including only the same subunits present in both the native PSII and Psb28–RC47), and Psb28–PSII and monomeric PSII (PDB: 3WU2) (including only the same subunits present in both the native PSII and Psb28–PSII) are 0.252 Å, 0.489 Å and 0.688 Å, respectively, indicating that the subunits shared by the Psb28–RC47 and Psb28–PSII complexes have similar structures (Extended Data Fig. 5) and are also similar to those of native PSII.

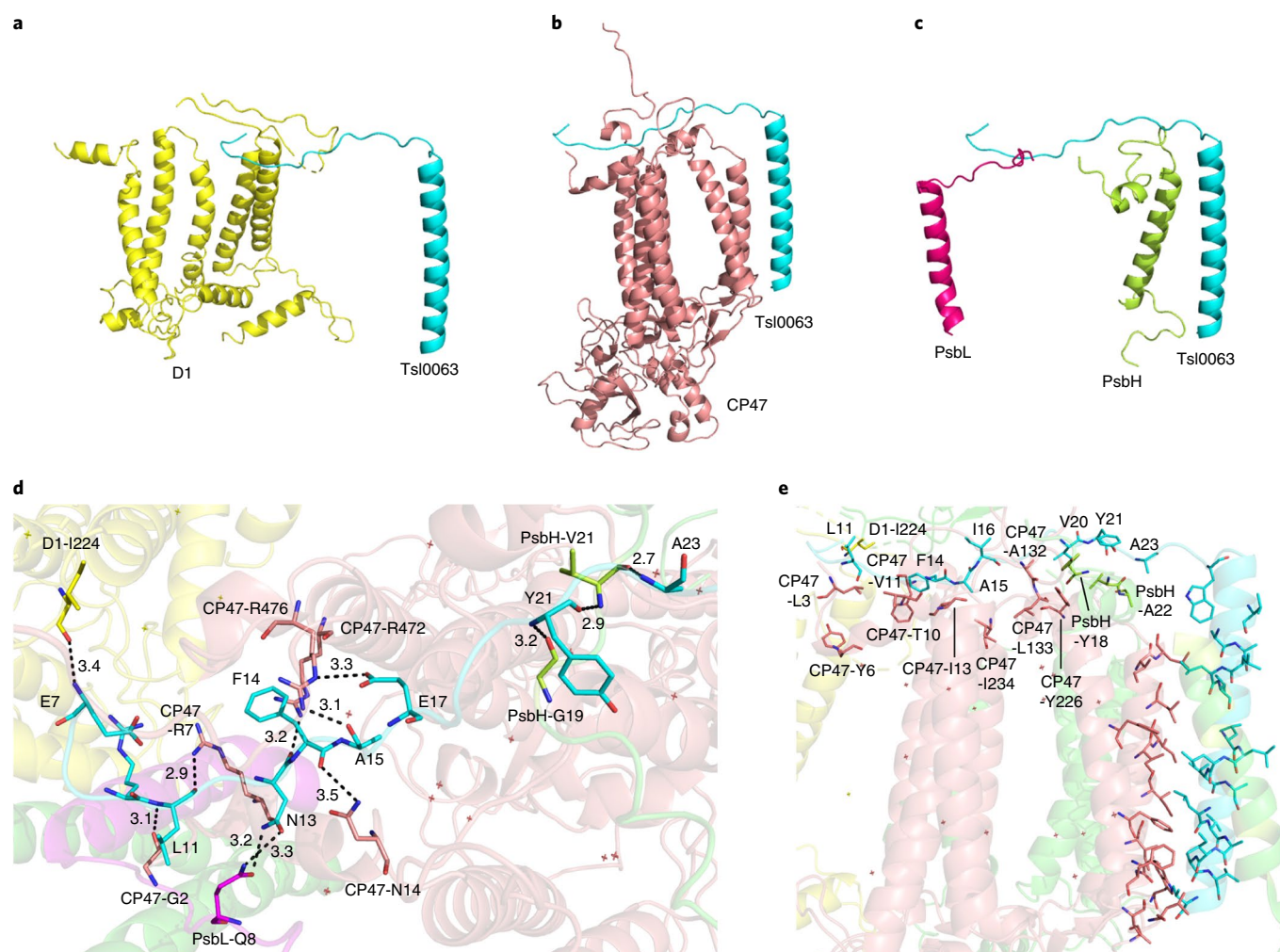
The extrinsic subunits PsbO, PsbV and PsbU present in the native PSII<sup>3,4</sup> are absent in both the Psb28–RC47 and Psb28–PSII complexes, consistent with the above SDS-PAGE analysis (Extended Data Figs. 1 and 2). PsbJ and PsbY were also not found in the cryo-EM structure, suggesting that they were lost during the sample purification owing to their loose association with the PSII core (Extended Data Fig. 1d). The loose association of PsbY with



**Fig. 3 | Structural comparison of the protein subunits between Psb28-PSII and native monomeric PSII. a, b**, Superposition of Psb28-PSII and native PSII from the side view (**a**) and top view from the stromal side (**b**). **c–i**, Superposition of D1 (**c**), D2 (**d**), CP47 (**e**), CP43 (**f**), PsbK (**g**), Psb30 (**h**) and PsbZ (**i**) subunits from Psb28-PSII and native PSII, showing the shifts and missing regions in subunits of the Psb28-PSII complex. The subunits of Psb28-PSII are coloured the same as those in Fig. 1, and the subunits of native PSII are in grey. C-ter, C terminus; N-ter, N terminus.

PSII has been shown in native PSII<sup>3,4</sup>, whereas the loss of PsbJ may be related to the assembly/repair process. In addition to the protein subunits, we identified 22 chlorophylls, 6 carotenoids, 2 pheophytins, 1 haem, 1 chloride ion and a number of other cofactors and lipids in the Psb28-RC47 monomer, and we additionally found 13

chlorophylls, 3 carotenoids and 1 chloride ion in the Psb28-PSII monomer (Extended Data Fig. 3 and Supplementary Table 3). No electron density corresponding to the  $\text{Mn}_4\text{CaO}_5$  cluster was observed in these two complexes, indicating the premature nature of the intermediate PSII.



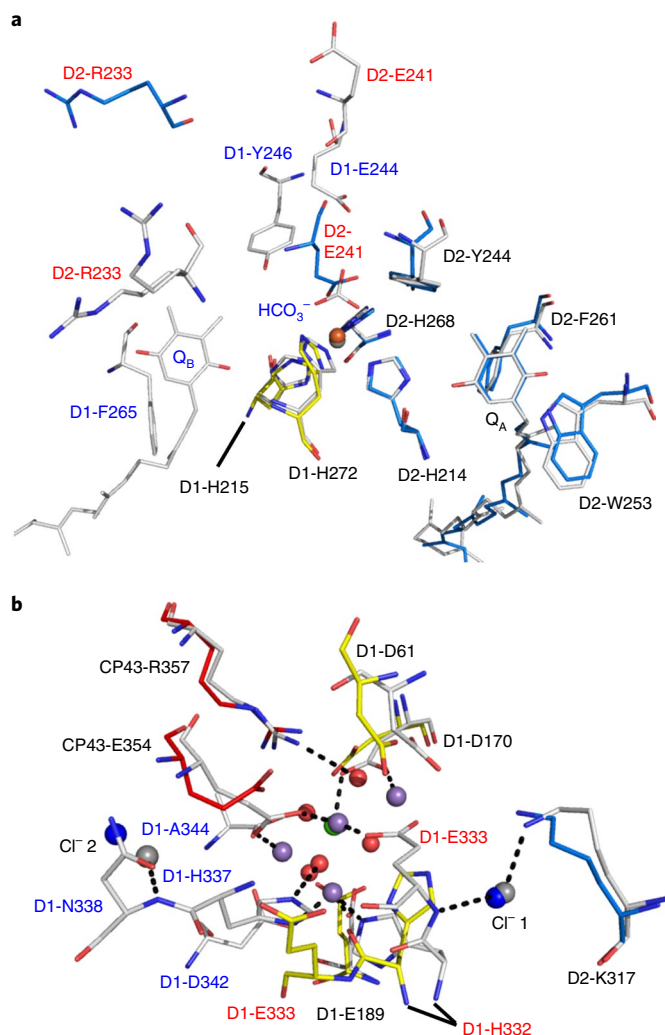
**Fig. 4 | Interactions between Tsl0063 and the PSII core in Psb28-RC47.** **a–c**, Relative positions of Tsl0063 and other subunits involved in the interactions with Tsl0063 in Psb28-RC47. The colours of the subunits in Psb28-RC47 are the same as those in Fig. 1. **d**, Hydrogen bonds formed between the amino acid residues of Tsl0063 and those of D1, CP47 and PsbH. The hydrogen bonds are shown in dashed lines, and their lengths are indicated in Å. **e**, Hydrophobic interactions between Tsl0063 and the CP47 and PsbH subunits. The amino acid residues involved in the interactions are represented as sticks and coloured yellow in D1, salmon in CP47, cyan in Tsl0063, lemon in PsbH and hot pink in PsbL.

**Interactions between Psb28 and the PSII core.** The cryo-EM structure of Psb28 in both Psb28-RC47 and Psb28-PSII is composed of seven  $\beta$ -strands, one long  $\alpha$ -helix and one short  $\alpha$ -helix, which is similar to the reported solution and crystal structures of recombinant Psb28 from cyanobacteria *Synechocystis* sp. PCC 6803 and *T. elongatus*, solved by NMR and X-ray crystallography, respectively<sup>40,41</sup> (Extended Data Fig. 6). In the cryo-EM structure of Psb28, the loop regions between  $\beta$ 2 and  $\beta$ 3 and between  $\beta$ 6 and  $\beta$ 7 are slightly shifted, and the C-terminus is located in a different direction from that in the crystal structure of recombinant *T. elongatus* Psb28 (Extended Data Fig. 6). These structural changes may be important in the interactions of Psb28 with the PSII core in the assembly intermediates.

The interactions between Psb28 and the PSII core were analysed in the Psb28-RC47 complex (Fig. 2). Psb28 binds to the PSII core by associating with D1, D2 and CP47 directly at the cytoplasmic surface of PSII (Fig. 2a), which is partly consistent with the previous predictions obtained by biochemical analysis<sup>28</sup>. The main interactions between Psb28 and D1, D2 and CP47 are found between Psb28 and the helix regions of the D1 D-E loop, the D2 D-E loop and the CP47 C-terminus. A large number of hydrogen bonds are found between the residues of Psb28 and those of the D1, D2 and

CP47 subunits (Fig. 2b,d and Table 1). Furthermore, there are two regions that form a number of hydrophobic interactions between residues of Psb28 and D1, D2 and CP47 (Fig. 2c,e), which may contribute to the stable binding of Psb28 in the PSII intermediate. One of these regions consists of the residues D1-Leu223 to D1-Ile224, D1-Ile248 to D1-Ala251, D1-Tyr254 to D1-Phe255, D2-Phe232 and D2-Ala234 to D2-Phe235, together with Psb28-Ile43 to Psb28-Val44, Psb28-Gly47, Psb28-Val51, Psb28-Val68, Psb28-Phe72, Psb28-Ile73 and Psb28-Ile80. The other region is formed by the C-terminal residues of CP47 (CP47-Val491, CP47-Trp493 and CP47-Phe495) and the N-terminal and C-terminal residues of Psb28 (Psb28-Val21, Psb28-Leu23, Psb28-Ala26, Psb28-Trp92, Psb28-Ile96, Psb28-Met99, Psb28-Ala103, Psb28-Leu108 and Psb28-Phe110). In spite of these interactions and the hydrogen bonds formed between residues Asn247, Ala250 and His252 of D1 and residues Asn74, Phe72 and Ala70 of Psb28, the regions before and after this binding area (D1-Glu231–Tyr246 and D1-Leu258–Ser270) are more flexible because part of the helices in these regions in the native PSII are changed to loops in Psb28-RC47 and Psb28-PSII, which results in a lower resolution in the cryo-EM density map (Fig. 3c).

Similar interactions of Psb28 with the PSII intrinsic subunits were observed in Psb28-PSII, and there were no interactions



**Fig. 5 | Structural changes around the non-haem iron and the  $\text{Mn}_4\text{CaO}_5$  cluster between Psb28-PSII and native PSII.** **a, b**, Superposition of the subunits and cofactors at the acceptor side (**a**) and donor side (**b**) between Psb28-PSII and native PSII (PDB ID code: 3WU2). The amino acid residues in native PSII are coloured in grey, and those of Psb28-PSII are coloured as in Fig. 1. Residues labelled in blue have a poor electron density in Psb28-PSII, and residues labelled in red have substantial structural changes in Psb28-PSII compared with their counterparts in native PSII.

between Psb28 and CP43, PsbK, Psb30 or PsbZ (Extended Data Fig. 5). These results suggest that Psb28 strongly binds to the PSII core in the RC47 assembly intermediate, which leads to a conformational change of the PSII core during PSII assembly or repair.

**Interactions between Tsl0063 and the PSII core.** Two subunits, Tsl0063 and UTP, are located close to CP47 and PsbH of Psb28-RC47 and Psb28-PSII in an opposite orientation with each other (Fig. 1). Tsl0063 has a very long N-terminal loop at the stromal side and interacts with D1, CP47, PsbH and PsbL through the N-terminal loop (Fig. 4). Hydrogen bonds are found between the residues of Tsl0063 and the D1, CP47, PsbL and PsbH subunits (Fig. 4d and Table 1). In addition, hydrophobic interactions are found between residues of the N-terminal loop of Tsl0063 and residues of PsbH and CP47, which may also contribute to the binding of Tsl0063 to the PSII core (Fig. 4e). UTP is located near Tsl0063 and may interact with PsbH and CP47; however, their exact interaction

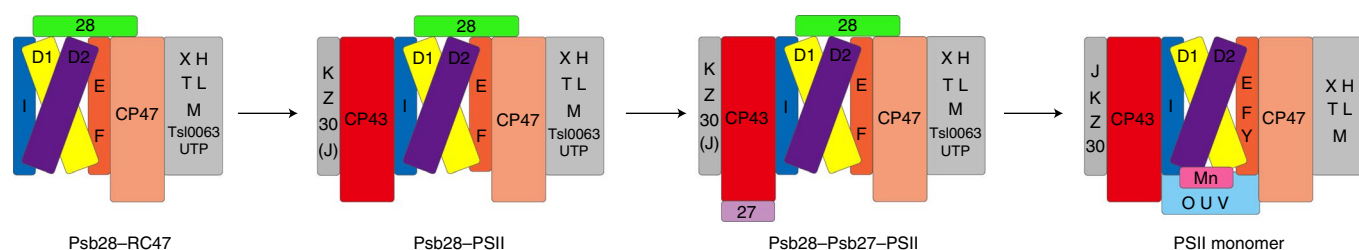
sites could not be determined, as the residues of UTP were assigned as alanines and not identified due to the limited resolution.

**Differences in the structures of Psb28-PSII and native PSII.** Since the Psb28-RC47 complex has a similar structure as the RC47 module of the Psb28-PSII complex (root mean square deviation value, 0.252 Å) (Extended Data Fig. 5), we compared the structure of Psb28-PSII including the CP43 module with that of native PSII to show a complete picture of the structural differences (Fig. 3). As expected, the overall structure of the core intrinsic subunits is similar between Psb28-PSII and native PSII. However, the monomeric Psb28-PSII exhibits some distinct differences in the structures of parts of the D1, D2, CP43, CP47, PsbK, Psb30 and PsbZ subunits compared with those of native PSII, as described below (Fig. 3c–i and Extended Data Fig. 7).

First, the D–E loop regions of D1 and D2 are important to maintaining the structural conformation at the stromal side in the native PSII. In particular, residues D1-His215, D1-Glu244, D1-Tyr246, D1-Phe265, D1-His272, D2-His214, D2-Glu241, D2-Tyr244, D2-Trp253, D2-Phe261 and D2-His268 are involved in the binding of the cofactors  $\text{Q}_\text{A}$ ,  $\text{Q}_\text{B}$ , bicarbonate ion and non-haem iron<sup>3,4</sup>. However, the structures of the D–E loop regions of D1 (Arg225–Arg269) and D2 (Asp225–Glu242) are shifted relative to their positions in native PSII due to the binding of Psb28 in the Psb28-PSII intermediate complex (Extended Data Fig. 8). This leads to conformational changes at the  $\text{Q}_\text{A}$  and  $\text{Q}_\text{B}$  binding sites as well as alterations in the coordination and the hydrogen-bond networks around the non-haem iron (Fig. 5a). Specifically, D1-Phe265 and the bicarbonate ion become invisible, and the original site of the bicarbonate is replaced by the residue D2-Glu241 at the acceptor side of Psb28-PSII. Two residues, D1-Glu244 and D1-Tyr246, that ligate the bicarbonate ion in the native PSII<sup>3,4</sup> also became difficult to recognize in the cryo-EM density map, probably due to the increased flexibility of these regions.

Second, in the PSII donor side, the structure around the  $\text{Mn}_4\text{CaO}_5$  cluster of Psb28-PSII exhibits larger differences than that of native PSII (Fig. 5b). The D1 C-terminal region, which contains the residues coordinated to the  $\text{Mn}_4\text{CaO}_5$  cluster in the native PSII<sup>3,4</sup>, shifts away from the cluster, and the region from Arg334 was not visible in Psb28-PSII, probably due to its higher flexibility in the absence of the  $\text{Mn}_4\text{CaO}_5$  cluster. Thus, among the seven residues that coordinate the  $\text{Mn}_4\text{CaO}_5$  cluster in the native PSII<sup>3,4</sup>, two residues, D1-Asp342 and D1-Ala344, could not be modelled due to the lack of apparent electron density. D1-Glu189 and D1-Asp170 are in similar locations as those in the native PSII, whereas dramatic changes in the positions of D1-His332 and D1-Glu333 are found in the Psb28-PSII intermediate (Fig. 5b). The residue D1-His332 shifts away from the  $\text{Mn}_4\text{CaO}_5$  cluster and interacts with the nearby  $\text{Cl}^-$  1 ion, and D1-Glu333 moves towards the luminal side. This alters the coordination of  $\text{Cl}^-$  1 compared with that in the native PSII. CP43-Glu354 is also shifted away from the  $\text{Mn}_4\text{CaO}_5$  cluster together with the rest of the soluble helix domain of CP43. In addition to the residues that ligate the  $\text{Mn}_4\text{CaO}_5$  cluster, residues D1-Asp61, D2-Lys317 and CP43-Arg357, as well as the  $\text{Cl}^-$  2 ion, also shifted slightly relative to their positions in the native PSII. This may reflect the structural conformation of the PSII donor side before the formation of the mature PSII.

Third, the regions of D1-Asn12 to D1-Arg16, D1-Asn298 to D1-Asn301, D1-Ala309 to D1-Gly311, CP47-Arg384 to CP47-Ser388 and CP47-Ser480 to CP47-Phe495 are shifted relative to their positions in the native PSII (Fig. 3c,e). The C-terminal residue Ala351 of D2 turns towards the luminal side (Fig. 3d). The loop regions of Glu231 to Thr246 and Leu258 to Ser270 of D1 are not resolved in the cryo-EM density map. All these changes may be correlated with the binding of Psb28 and Tsl0063 and the loss of extrinsic subunits at the luminal side in the PSII assembly intermediate.



**Fig. 6 | A model proposed for the assembly of PSII.** The model emphasizes the assembly of CP43 including PsbK, Psb30 and PsbZ, on the basis of the structural information of two PSII intermediate complexes, Psb28-RC47 and Psb28-PSII, solved in the present study. The components are shown in the following colours: D1, yellow; D2, purple; CP43, red; CP47, light orange; and Psb28, green. The subunits PsbE, PsbF, PsbH, PsbI, PsbJ, PsbK, PsbL, PsbM, PsbT, PsbX, PsbY, PsbZ, PsbO, PsbU, PsbV, Psb27, Psb28 and Psb30 are shown as E, F, H, I, J, K, L, M, T, X, Y, Z, O, U, V, Psb27, Psb28 and Psb30, respectively, and are labelled with different colours in the figure. Mn represents the Mn cluster. See the text for more details.

In addition, slight shifts of helices are observed in the CP43, PsbK, Psb30 and PsbZ subunits (Fig. 3f–i) of the Psb28-PSII complex, indicating that these subunits may also be involved in the assembly process of PSII.

## Discussion

Psb28-RC47 and Psb28-PSII are PSII assembly intermediates that lack the  $\text{Mn}_4\text{CaO}_5$  cluster and three extrinsic proteins and do not have any oxygen-evolving activity. The cryo-EM structures of these two complexes from *T. vulcanus* showed the binding of Psb28 and Tsl0063, two cofactors required for the assembly of mature PSII. The cofactor Psb28 has been reported previously, whereas the cofactor Tsl0063 has been found in the present study and in a very recent study<sup>19</sup>. While one copy of Tsl0063 was found in the previous study<sup>19</sup>, we found one copy of Tsl0063 and one copy of an unidentified transmembrane protein subunit in the immediate vicinity of Tsl0063. The exact sequences of the second copy could not be identified unambiguously, so this second copy may be another Tsl0063, another homologous subunit in the CAB/ELIP/HLIP protein family or a homologous subunit of a recently reported cyanobacterial Psb35 (ref. 43). Tsl0063 was named Psb34 by Zabret et al.<sup>19</sup>; however, the name Psb34 had already been used for a PSII transmembrane subunit found in diatom PSII<sup>7</sup>. Given that the name Psb35 has recently been used for another subunit<sup>43</sup>, another name such as Psb36 should be used for Tsl0063.

Psb28-RC47 lacks CP43, PsbK, Psb30 and PsbZ, whereas Psb28-PSII has these four subunits attached. This supports the view that D1, D2 and CP47 first assemble into a pre-complex, followed by the assembly of CP43 and several other small subunits<sup>11–14</sup>. While Psb28-RC47 has a similar structural arrangement as that of Psb28-PSII, the additional CP43, PsbK, Psb30 and PsbZ subunits in Psb28-PSII are shifted away from D1 in comparison with their counterparts in the native PSII. This indicates the loose binding of these subunits in Psb28-PSII as well as the necessity of their structural changes during the full assembly of the native PSII.

Psb28 was found to be anchored on the cytoplasmic surface of PSII by interacting with D1, D2 and CP47. In particular, the region from Asn247 to Gly253 of D1, the region from Phe232 to Asn236 of D2 and the C-terminal region of CP47 were found to play important roles in the binding of Psb28 to PSII. These interactions are largely similar to the results reported very recently<sup>19</sup>. However, in that report, Psb27 (another assembly cofactor of PSII) was associated at the luminal side, which is absent in the present PSII assembly intermediate. Although Psb27-PSII was also found in part of the  $\Delta\text{PsbV}$  strain in the present and previous studies<sup>18,42</sup>, it does not bind to the PSII complex to which Psb28 was bound. This indicates that Psb28 and Psb27 may aid the assembly of PSII intermediates in different processes.

Tsl0063 has a high sequence similarity to CAB/ELIP/HLIP family proteins of cyanobacteria. The CAB/ELIP/HLIP family

proteins have been shown to support numerous functions such as the regulation of chlorophyll biosynthesis<sup>44,45</sup>, the transient binding of chlorophylls and carotenoids<sup>46–48</sup>, chlorophyll turnover<sup>49</sup>, non-photochemical energy quenching<sup>50</sup>, and the prevention of the formation of singlet oxygen<sup>51,52</sup>. Most of these functions are related to pigments (chlorophylls or carotenoids) during pigment–protein synthesis and assembly<sup>39</sup>. In contrast, Tsl0063 does not bind chlorophylls; this suggests that it may be involved in other functions that are independent of pigment biosynthesis and binding during the assembly of PSII. Indeed, Tsl0063 has strong interactions with CP47 as well as D1, PsbH and PsbL in the stromal side of PSII. Tsl0063 may thus play an important role in stabilizing the CP47-containing complex or maintaining a structural conformation that enables the binding of Psb28 during the PSII assembly cycle.

Large structural changes of the D–E region (Arg225–Arg269) of D1 and the D–E region (Asp225–Glu242) of D2 are observed in the Psb28-PSII intermediate compared with the native PSII, due to the binding of Psb28. This leads to conformational changes at the  $Q_A$  and  $Q_B$  binding sites as well as alterations of coordination and the hydrogen-bond networks around the non-haem iron. Such structural changes may not only slow down the electron transfer to the plastoquinone pool but also change the potential of  $Q_A/Q_A^{\bullet-}$  to a more positive value, as suggested previously<sup>53</sup>. This positive shift is predicted to increase the energy gap between  $P_{680}^{+/+}/Q_A^{\bullet-}$  and  $P_{680}^{+/+}/\text{Pheo}^{\bullet-}$ , which may decrease the back-reactions via the  $P_{680}^{+/+}/\text{Pheo}^{\bullet-}$  recombination route and reduce the formation of the RC triplet state<sup>53</sup>. This may lead to a decrease in the production of  $^1\text{O}_2$ , a product causing photodamage. The Psb28- and Tsl0063-induced conformational changes may therefore benefit the protection of the premature PSII from photodamage.

The structure of the donor side of Psb28-PSII represents the  $\text{Mn}_4\text{CaO}_5$  binding site before its binding. Large structural changes of the donor side were found in the Psb28-PSII intermediate compared with that of the native PSII; these changes were similar to those of a recently reported apo-monomeric PSII from *Synechocystis* PCC 6803 (ref. 16). The disorder of the coordinating residues of the  $\text{Mn}_4\text{CaO}_5$  cluster observed in Psb28-PSII is also in line with the structure observed for Psb27-PSII<sup>18,19,54</sup>. In addition, the C-terminal residue Ala351 of D2 turns towards the luminal side, resulting in the breakage of the interactions between D2 and PsbU seen in the native PSII, which may be one of the reasons for the inability of PsbU to bind to PSII. The C-terminal helix of CP47 in native PSII is changed to a loop in the Psb28-RC47 and Psb28-PSII intermediates due to interactions with the Psb28 subunit, whereas structural changes in the region of CP47-Arg384 to CP47-Ser388 remain unknown.

On the basis of the above results, we propose a schematic model for the assembly of PSII (Fig. 6), combining the current structural information with the results of previous studies<sup>12,14,27,34,35,54,55</sup>. In the

early steps of PSII assembly, one Psb28 and one Tsl0063 subunit bind to PSII to form the RC47 complex, which has a protein composition as shown in the present cryo-EM structure. Subsequently, the CP43 module consisting of CP43, PsbK, Psb30 and PsbZ is attached, and then Psb27 binds to the CP43 module at the luminal side, leading to the formation of the monomeric Psb28–Psb27–Tsl0063–PSII complex. Finally, Psb28 and Tsl0063 are detached, followed by the assembly of the  $\text{Mn}_4\text{CaO}_5$  cluster, the dissociation of Psb27 and the insertion of PsbJ as well as the attachment of the three extrinsic proteins to form an active dimer. However, it is not clear how and when Psb28, Tsl0063 and Psb27 are released during the assembly. It has been suggested that the conformational rearrangement of the luminal domain of CP43 plays an important role in the dissociation of Psb27 (refs. <sup>14,55</sup>), and the incorporation of the  $\text{Mn}_4\text{CaO}_5$  cluster might trigger this dynamic rearrangement. It was proposed that the Psb28 dissociation occurs before the binding of Psb27 on the basis of the result that the purified, His-tagged Psb27–PSII complex lacks Psb28 (ref. <sup>35</sup>), but how Psb28 and Tsl0063 are detached from the assembly intermediate remains elusive. Strong interactions between Psb28 and PSII are found in the assembly intermediate, which means that PSII needs to undergo substantial structural changes to dissociate from Psb28. Finally, the identity and role of the unidentified subunit in the assembly of PSII remain to be clarified.

In conclusion, both structures of Psb28–RC47 and Psb28–PSII contain one Psb28 and one Tsl0063 subunit, as well as one UTP. Psb28 mainly interacts with D1, D2 and CP47, whereas Tsl0063 interacts with D1, CP47, PsbH and PsbL. Our results reveal the structural and functional relationships between the assembly factors Psb28, Tsl0063 and the PSII core and provide important structural insights into the assembly process of native PSII.

## Methods

**Construction of His-tagged *psb28* strains and culture.** To facilitate the purification of PSII assembly intermediates associated with Psb28, a hexa-histidine tag was attached to the C-terminus of the Psb28 protein. To generate the His-tagged Psb28 strain of *T. vulcanus*, a His-*psb28*-pUC18 plasmid vector, which includes a His<sub>6</sub> tag at the downstream of *psb28*, a chloramphenicol-resistant gene cassette and the upstream and downstream of *psb28* for homologous recombination, was constructed as follows (Supplementary Fig. 1a). First, a 5'-flanking region containing the *psb28* gene fragment was amplified from the genome of *T. vulcanus* by PCR using the F-1 and R-1 primers (Supplementary Table 1). Second, using this 5'-flanking region as the template, a PCR fragment containing a 5'-flanking region, the *psb28* gene and the His<sub>6</sub> tag was amplified using the F-1 and R-His (containing the sequence of the His<sub>6</sub> tag) primers (the 5'-flanking region containing the *psb28*-His-tag fragment). Third, the 3'-flanking fragment of the *psb28* gene was amplified from the genome of *T. vulcanus* using F-2 and R-2 as primers, and a chloramphenicol-resistant gene cassette was amplified from the CDFDuet-1 plasmid using F-3 and R-3 as primers (Supplementary Table 1). All the PCR fragments obtained were purified by a DNA gel extraction kit (Axygen) and cloned into the pUC18 plasmid with the chloramphenicol-resistant gene cassette located between the His tag and the 3'-flanking fragment. This plasmid was named His-*psb28*-pUC18 and introduced into the WT and *psbV*-deleted mutant ( $\Delta\text{PsbV}$ )<sup>42</sup> cells of *T. vulcanus* by natural transformation as described previously<sup>56,57</sup>. After the appearance of the transformants in two to five weeks, the clones were segregated by growing on agar plates containing  $5\text{ }\mu\text{g ml}^{-1}$  chloramphenicol at least three times.

To confirm the transformed *T. vulcanus* cells, DNA was amplified by PCR with primers F and R, and the products obtained were analysed by electrophoresis as well as DNA sequencing. The WT,  $\Delta\text{PsbV}$ , HisPsb28-WT and HisPsb28- $\Delta\text{PsbV}$  cells were grown in a medium described in Xiao et al.<sup>42</sup> and bubbled with air containing 3–5% (v/v)  $\text{CO}_2$  at 45°C under continuous light.

**Preparation of PSII core and assembly intermediate complexes.** The PSII core complexes from WT cells were prepared using the protocol of Shen et al.<sup>58</sup>. His-tagged PSII assembly intermediate complexes were purified from HisPsb28-WT or HisPsb28- $\Delta\text{PsbV}$  strains using a Ni-affinity column. The thylakoids were solubilized by 1%  $\beta$ -DDM (w/v), and the solubilized membranes were loaded onto a Ni-Sepharose HP column, followed by elution as described in Sugiura and Inoue<sup>59</sup>. The Psb28-containing PSII assembly intermediate was further purified by a MonoQ anion exchange column, which had been equilibrated with 30 mM HEPES-NaOH (pH 8.0), 3 mM  $\text{CaCl}_2$ , 0.03%  $\beta$ -DDM (w/v) and 5% glycerol (w/v), followed by elution with a linear gradient of NaCl. The eluted PSII intermediates were concentrated to a chlorophyll concentration above 1 mg chlorophyll per ml

using a 100 kDa cut-off membrane concentrator, and they were stored in 40 mM MES-NaOH (pH 6.5), 10 mM NaCl, 10 mM  $\text{CaCl}_2$ , 10 mM  $\text{MgCl}_2$ , 0.03%  $\beta$ -DDM (w/v) and 25% glycerol (w/v) until use.

**Biochemical analysis.** The protein composition of the purified PSII assembly intermediate was analysed by SDS–PAGE with a 16% polyacrylamide gel containing 7.5 M urea<sup>60</sup>. For analysis of the PSII dimer and monomer, BN–PAGE was performed according to Kawakami et al.<sup>61</sup>. For analysis of the Tsl0063 and PsbJ subunits, Tricine SDS–PAGE was used according to Schägger<sup>62</sup>. For immunoblotting, proteins on the gel were transferred to a polyvinylidene fluoride membrane and incubated with the antibodies against the respective proteins. The subunit bands were visualized by enhanced chemiluminescence (Biopstep) after incubation with a horseradish-peroxidase-conjugated secondary antibody. The primary antibody against His-tag (A02051) was purchased from Abbkine, and the primary antibodies against D1 (AS05084), D2 (AS06146), PsbO (AS06142–33) and the horseradish-peroxidase-conjugated secondary antibody (goat anti-rabbit IgG) (AS09602) were purchased from Agrisera. The primary antibodies against Psb27, PsbJ and Tsl0063 were custom-made by Genscript.

Mass spectrometry (MS) analysis was performed as described in Shen et al.<sup>9</sup>. Targeted protein bands were excised from Coomassie-brilliant-blue-stained SDS–PAGE gel and digested using sequencing grade-modified trypsin in 50 mM ammonium bicarbonate at 37°C. The peptides were extracted twice with 1% trifluoroacetic acid in 50% acetonitrile aqueous solution for 30 min. For liquid chromatography with tandem MS (LC–MS/MS) analysis, the peptides were separated using a 120 min gradient elution at a flow rate of  $0.30\text{ }\mu\text{l min}^{-1}$  with a Thermo-Dionex Ultimate 3000 HPLC system, which was directly interfaced with the Thermo Orbitrap Fusion mass spectrometer. The analytical column was a fused silica capillary column (75  $\mu\text{m}$  inner diameter, 150 mm length; Upchurch) packed with C-18 resin. The MS/MS spectra from each LC–MS/MS run were searched against the selected database using the Proteome Discovery searching algorithm (version 1.4).

**Cryo-EM sample preparation and data collection.** In the mixture samples of Psb28–PSII and Psb28–RC47, the samples with a high proportion of Psb28–PSII and a high proportion of Psb28–RC47 were used for the collection of the Psb28–PSII and Psb28–RC47 cryo-EM micrographs, respectively. A droplet of 4  $\mu\text{l}$  of the Psb28–PSII and Psb28–RC47 mixture sample at a concentration of around 1.4 mg chlorophyll per ml was applied to a glow-discharged Quantifoil holey carbon Cu grid (R 1.2/1.3, 400 mesh), and plunged into liquid ethane at around 100 K using a FEI Vitrobot Mark IV. The parameters were set as follows: blotting time, 5 s to 6 s; blotting force level, 0; humidity, 100%; and temperature, 8°C. Sample screening was performed using a Tecnai Arctica 200 keV electron microscope equipped with a FEI Falcon II camera. Cryo-EM data used for high-resolution structure determination were collected with a 300 keV FEI Titan Krios electron microscope equipped with a K2 Summit direct electron detector (Gatan) in the super-resolution mode, using SerialEM v.3.8. A total of 2,207 micrographs for the Psb28–PSII dataset and 3,461 micrographs for the Psb28–RC47 dataset were recorded at a nominal magnification of  $\times 22,500$ , yielding a pixel size of  $1.30654\text{ }\text{\AA}$ , with defocus values ranging from  $-1.5$  to  $-2.5\text{ }\mu\text{m}$ . Each exposure of 8 s was dose-fractionated to 32 video frames, leading to a total dose of approximately 50 electrons per  $\text{\AA}^2$ .

**Cryo-EM image processing.** For the data collected on the Titan Krios, the beam-induced motion of the whole micrograph with 32 video frames was corrected by MotionCorr2 (ref. <sup>63</sup>), and contrast transfer function parameters were estimated by CTFIND4 (ref. <sup>64</sup>) using non-dose-weighted micrographs. Each micrograph was manually inspected to remove images that contained crystalline ice or contamination or that were taken from carbon films; then, 1,764 micrographs and 2,931 micrographs were selected for subsequent Psb28–PSII and Psb28–RC47 data processing, respectively.

For the structural analysis of Psb28–PSII, a small dataset of around 2,000 particles were manually picked and processed by reference-free 2D classification using RELION v.3.0 (ref. <sup>65</sup>). Nine good 2D averages were selected as references for auto-picking of all 1,764 micrographs, and a total of 1,442,584 particles were automatically picked using RELION 3.0. To speed up calculation, the whole dataset was evenly divided into two subsets and then subjected to the reference-free 2D classification. In total, 1,416,275 particles were selected from the two subsets and subjected to 3D classification. A density map was generated from an atomic model of monomer PSII core (PDB code: 3WU2) without the three extrinsic proteins and was low-pass filtered to 60  $\text{\AA}$  as the initial model for the 3D classification. Two rounds of 3D classification with global search at  $2.61308\text{ }\text{\AA}$  per pixel were used for discarding junk particles, and another 3D classification without image alignment at  $1.30654\text{ }\text{\AA}$  per pixel was used for discarding particles without Psb28 and/or CP43. Finally, 194,738 homogeneous Psb28–PSII particles were selected for 3D auto-refinement.

For the structural analysis of Psb28–RC47, a total of 3,759,725 particles were auto-picked using a template-free procedure based on a Laplacian-of-Gaussian filter and extracted in 200-by-200-pixel boxes using RELION v.3.0. The whole dataset was divided into four subsets and subjected to reference-free 2D

classification. Poorly defined classes were discarded, resulting in 3,467,418 particles for further 3D processing. The initial model was derived from the previous data processing results of Psb28–PSII. Two rounds of 3D classification at 2.61308 Å per pixel and another 3D classification at 1.30654 Å per pixel were performed to discard bad particles, which resulted in a final set of 241,790 homogeneous Psb28–RC47 particles for the subsequent 3D auto-refinement.

Further quality improvement of the density maps was achieved by re-extracting good particles from dose-weighted micrographs and conducting 3D auto-refinement and contrast transfer function refinement in repeated cycles. The final resolution of both Psb28–PSII and Psb28–RC47 density maps was 3.14 Å on the basis of the gold-standard FSC = 0.143 criteria<sup>66</sup>. Local resolution estimates were determined with RELION v3.1.

**Model building and refinement.** The crystal structure of monomer PSII core from *T. vulcanus* PSII (PDB code: 3WU2) with the CP43, PsbJ, PsbK, PsbO, PsbU, PsbV, Psb30 and PsbZ subunits manually deleted and the 2.4 Å resolution X-ray structure of Psb28 from *T. vulcanus* (PDB code: 3ZPN) were manually fitted into the Psb28–RC47 cryo-EM map using UCSF Chimera<sup>67</sup>. Regarding the two subunits found in both the Psb28–RC47 and Psb28–PSII complexes, the first subunit has an obvious membrane-spanning  $\alpha$ -helix and a long loop in the cytoplasmic side. Combining this with the results of mass spectra, secondary structure prediction, transmembrane region prediction and western blotting, we were able to assign the density to Tsl0063 and modelled most of the side chains of Tsl0063 to the cryo-EM density using Coot<sup>68</sup>. The cryo-EM density of the other subunit was not clear enough to allow the assignment of side chains; thus, we were not able to identify the subunit and built the model with poly-alanines.

We then used Coot to correct and adjust the positions of side chains manually on the basis of the cryo-EM map. We also deleted some residues where their cryo-EM density maps were not clear. Next, we ran several rounds of refinement with phenix.real\_space\_refine<sup>69</sup>. The refined model was corrected again in Coot until there were no more improvements in both MolProbity score and geometry parameters.

The structure of Psb28–PSII was built with the model of Psb28–RC47 merged with the CP43, PsbK, Psb30 and PsbZ subunits from *T. vulcanus* PSII (PDB code: 3WU2) as the initial model for Psb28–PSII. The final atomic model of Psb28–PSII was completed via iterative rounds of manual refinement with Coot and auto-refinement with phenix.real\_space\_refine. The final statistics for data processing and structure refinement are summarized in Supplementary Table 2.

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

## Data availability

The cryo-EM density maps and atomic models of the Psb28–RC47 and Psb28–PSII complexes at 3.14 Å resolution have been deposited in the Electron Microscopy Data Bank and the Protein Data Bank (EMD ID 30902 and PDB ID 7DXA for Psb28–RC47, EMD ID 30909 and PDB ID 7DXH for Psb28–PSII). The data that support the findings of this study are available from the corresponding authors upon reasonable request. Source data are provided with this paper.

Received: 20 January 2021; Accepted: 3 June 2021;  
Published online: 5 July 2021

## References

- Barber, J. Photosystem II: the engine of life. *Q. Rev. Biophys.* **36**, 71–89 (2003).
- Shen, J.-R. The structure of photosystem II and the mechanism of water oxidation in photosynthesis. *Annu. Rev. Plant Biol.* **66**, 23–48 (2015).
- Umena, Y., Kawakami, K., Shen, J. R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55–60 (2011).
- Suga, M. et al. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99–103 (2015).
- Ago, H. et al. Novel features of eukaryotic photosystem II revealed by its crystal structure analysis from a red alga. *J. Biol. Chem.* **291**, 5676–5687 (2016).
- Wei, X. et al. Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. *Nature* **534**, 69–74 (2016).
- Pi, X. et al. The pigment–protein network of a diatom photosystem II–light-harvesting antenna supercomplex. *Science* **365**, eaax4406 (2019).
- Nagao, R. et al. Structural basis for energy harvesting and dissipation in a diatom PSII–FCPII supercomplex. *Nat. Plants* **5**, 890–901 (2019).
- Shen, L. et al. Structure of a C2S2M2N2-type PSII–LHCII supercomplex from the green alga *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA* **116**, 21246–21255 (2019).
- Sheng, X. et al. Structural insight into light harvesting for photosystem II in green algae. *Nat. Plants* **5**, 1320–1330 (2019).
- Nixon, P. J., Michoux, F., Yu, J., Boehm, M. & Komenda, J. Recent advances in understanding the assembly and repair of photosystem II. *Ann. Bot.* **106**, 1–16 (2010).
- Boehm, M. et al. Investigating the early stages of photosystem II assembly in *Synechocystis* sp. PCC 6803: isolation of CP47 and CP43 complexes. *J. Biol. Chem.* **286**, 14812–14819 (2011).
- Boehm, M. et al. Subunit composition of CP43-less photosystem II complexes of *Synechocystis* sp. PCC 6803: implications for the assembly and repair of photosystem II. *Phil. Trans. R. Soc. B* **367**, 3444–3454 (2012).
- Komenda, J., Sobotka, R. & Nixon, P. J. Assembling and maintaining the photosystem II complex in chloroplasts and cyanobacteria. *Curr. Opin. Plant Biol.* **15**, 245–251 (2012).
- Zhang, M. et al. Structural insights into the light-driven auto-assembly process of the water-oxidizing Mn4CaO5-cluster in photosystem II. *eLife* **6**, e26933 (2017).
- Gisriel, C. J. et al. Cryo-EM structure of monomeric photosystem II from *Synechocystis* sp. PCC 6803 lacking the water-oxidation complex. *Joule* **4**, 2131–2148 (2020).
- Tokano, T., Kato, Y., Sugiyama, S., Uchihashi, T. & Noguchi, T. Structural dynamics of a protein domain relevant to the water-oxidizing complex in photosystem II as visualized by high-speed atomic force microscopy. *J. Phys. Chem. B* **124**, 5847–5857 (2020).
- Huang, G. et al. Structural insights into a dimeric Psb27–photosystem II complex from a cyanobacterium *Thermosynechococcus vulcanus*. *Proc. Natl Acad. Sci. USA* **118**, e2018053118 (2021).
- Zabret, J. et al. Structural insights into photosystem II assembly. *Nat. Plants* **7**, 524–538 (2021).
- Bao, H. & Burnap, R. L. Photoactivation: the light-driven assembly of the water oxidation complex of photosystem II. *Front. Plant Sci.* **7**, 578 (2016).
- Nickelsen, J. & Rengstl, B. Photosystem II assembly: from cyanobacteria to plants. *Annu. Rev. Plant Biol.* **64**, 609–635 (2013).
- Jarvi, S., Suorsa, M. & Aro, M. E. Photosystem II repair in plant chloroplasts—regulation, assisting proteins and shared components with photosystem II biogenesis. *Biochim. Biophys. Acta* **1847**, 900–909 (2015).
- Heinz, S., Liauw, P., Nickelsen, J. & Nowaczyk, M. Analysis of photosystem II biogenesis in cyanobacteria. *Biochim. Biophys. Acta* **1857**, 274–287 (2016).
- Kashino, Y. et al. Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. *Biochemistry* **41**, 8004–8012 (2002).
- Nowaczyk, M. M. et al. Psb27, a cyanobacterial lipoprotein, is involved in the repair cycle of photosystem II. *Plant Cell* **18**, 3121–3131 (2006).
- Avramov, A. P., Hwang, H. J. & Burnap, R. L. The role of Ca<sup>2+</sup> and protein scaffolding in the formation of nature's water oxidizing complex. *Proc. Natl Acad. Sci. USA* **117**, 28036–28045 (2020).
- Mabitt, P. D., Wilbanks, S. M. & Eaton-Rye, J. J. Structure and function of the hydrophilic Photosystem II assembly proteins: Psb27, Psb28 and Ycf48. *Plant Physiol. Biochem.* **81**, 96–107 (2014).
- Dobakova, M., Sobotka, R., Tichy, M. & Komenda, J. Psb28 protein is involved in the biogenesis of the photosystem II inner antenna CP47 (PsbB) in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol.* **149**, 1076–1086 (2009).
- Sakata, S., Mizusawa, N., Kubota-Kawai, H., Sakurai, I. & Wada, H. Psb28 is involved in recovery of photosystem II at high temperature in *Synechocystis* sp. PCC 6803. *Biochim. Biophys. Acta* **1827**, 50–59 (2013).
- Beckova, M. et al. Association of Psb28 and Psb27 proteins with PSII–PSI supercomplexes upon exposure of *Synechocystis* sp. PCC 6803 to high light. *Mol. Plant* **10**, 62–72 (2017).
- Jung, K. H. et al. Identification and functional analysis of light-responsive unique genes and gene family members in rice. *PLoS Genet.* **4**, e1000164 (2008).
- Yao, D. et al. Localization of the small CAB-like proteins in photosystem II. *J. Biol. Chem.* **282**, 267–276 (2007).
- Liu, H., Roose, J. L., Cameron, J. C. & Pakrasi, H. B. A genetically tagged Psb27 protein allows purification of two consecutive photosystem II (PSII) assembly intermediates in *Synechocystis* 6803, a cyanobacterium. *J. Biol. Chem.* **286**, 24865–24871 (2011).
- Nowaczyk, M. M. et al. Deletion of psbJ leads to accumulation of Psb27–Psb28 photosystem II complexes in *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta* **1817**, 1339–1345 (2012).
- Weisz, D. A. et al. Mass spectrometry-based cross-linking study shows that the Psb28 protein binds to cytochrome b559 in Photosystem II. *Proc. Natl Acad. Sci. USA* **114**, 2224–2229 (2017).
- Weisz, D. A. et al. A novel chlorophyll protein complex in the repair cycle of photosystem II. *Proc. Natl Acad. Sci. USA* **116**, 21907–21913 (2019).
- Funk, C. & Vermaas, W. A cyanobacterial gene family coding for single-helix proteins resembling part of the light-harvesting proteins from higher plants. *Biochemistry* **38**, 9397–9404 (1999).
- He, Q., Dolganov, N., Bjorkman, O. & Grossman, A. R. The high light-inducible polypeptides in *Synechocystis* PCC6803: expression and function in high light. *J. Biol. Chem.* **276**, 306–314 (2001).
- Komenda, J. & Sobotka, R. Cyanobacterial high-light-inducible proteins—protectors of chlorophyll-protein synthesis and assembly. *Biochim. Biophys. Acta* **1857**, 288–295 (2016).

40. Yang, Y. et al. Solution NMR structure of photosystem II reaction center protein Psb28 from *Synechocystis* sp. strain PCC 6803. *Proteins* **79**, 340–344 (2011).
41. Bialek, W. et al. Crystal structure of the Psb28 accessory factor of *Thermosynechococcus elongatus* photosystem II at 2.3 Å. *Photosynth. Res.* **117**, 375–383 (2013).
42. Xiao, Y. et al. Role of PsbV-Tyr137 in photosystem II studied by site-directed mutagenesis in the thermophilic cyanobacterium *Thermosynechococcus vulcanus*. *Photosynth. Res.* **146**, 41–54 (2020).
43. Pascual-Aznar, G. et al. Psb35 protein stabilizes the CP47 assembly module and associated high-light inducible proteins during the biogenesis of photosystem II in the cyanobacterium *Synechocystis* sp. PCC6803. *Plant Cell Physiol.* **61**, 178–190 (2021).
44. Xu, H., Vavilin, D., Funk, C. & Vermaas, W. Small Cab-like proteins regulating tetrapyrrole biosynthesis in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Mol. Biol.* **49**, 149–160 (2002).
45. Hernandez-Prieto, M. A. et al. The small CAB-like proteins of the cyanobacterium *Synechocystis* sp. PCC 6803: their involvement in chlorophyll biogenesis for Photosystem II. *Biochim. Biophys. Acta* **1807**, 1143–1151 (2011).
46. Xu, H., Vavilin, D., Funk, C. & Vermaas, W. Multiple deletions of small Cab-like proteins in the cyanobacterium *Synechocystis* sp. PCC 6803: consequences for pigment biosynthesis and accumulation. *J. Biol. Chem.* **279**, 27971–27979 (2004).
47. Knoppová, J. et al. Discovery of a chlorophyll binding protein complex involved in the early steps of photosystem II assembly in *Synechocystis*. *Plant Cell* **26**, 1200–1212 (2014).
48. Staleva, H. et al. Mechanism of photoprotection in the cyanobacterial ancestor of plant antenna proteins. *Nat. Chem. Biol.* **11**, 287–291 (2015).
49. Vavilin, D., Yao, D. & Vermaas, W. Small Cab-like proteins retard degradation of photosystem II-associated chlorophyll in *Synechocystis* sp. PCC 6803: kinetic analysis of pigment labeling with <sup>15</sup>N and <sup>13</sup>C. *J. Biol. Chem.* **282**, 37660–37668 (2007).
50. Havaux, M., Guedeney, G., He, Q. F. & Grossman, A. R. Elimination of high-light-inducible polypeptides related to eukaryotic chlorophyll *a/b*-binding proteins results in aberrant photoacclimation in *Synechocystis* PCC6803. *Biochim. Biophys. Acta* **1557**, 21–33 (2003).
51. Sinha, R. K., Komenda, J., Knoppová, J., Sedlarova, M. & Pospisil, P. Small CAB-like proteins prevent formation of singlet oxygen in the damaged photosystem II complex of the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Environ.* **35**, 806–818 (2012).
52. Tibiletti, T., Rehman, A. U., Vass, I. & Funk, C. The stress-induced SCP/HLIP family of small light harvesting-like proteins (ScpABCDE) protects Photosystem II from photoinhibitory damages in the cyanobacterium *Synechocystis* sp. PCC 6803. *Photosynth. Res.* **135**, 103–114 (2018).
53. Brinkert, K., De Causmaecker, S., Krieger-Liszka, A., Fantuzzi, A. & Rutherford, A. W. Bicarbonate-induced redox tuning in Photosystem II for regulation and protection. *Proc. Natl Acad. Sci. USA* **113**, 12144–12149 (2016).
54. Liu, H., Huang, R. Y., Chen, J., Gross, M. L. & Pakrasi, H. B. Psb27, a transiently associated protein, binds to the chlorophyll binding protein CP43 in photosystem II assembly intermediates. *Proc. Natl Acad. Sci. USA* **108**, 18536–18541 (2011).
55. Liu, H. et al. Mass spectrometry-based footprinting reveals structural dynamics of loop E of the chlorophyll-binding protein CP43 during photosystem II assembly in the cyanobacterium *Synechocystis* 6803. *J. Biol. Chem.* **288**, 14212–14220 (2013).
56. Muhlenhoff, U. & Chauvat, F. Gene transfer and manipulation in the thermophilic cyanobacterium *Synechococcus elongatus*. *Mol. Gen. Genet.* **252**, 93–100 (1996).
57. Onai, K., Morishita, M., Kaneko, T., Tabata, S. & Ishiura, M. Natural transformation of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1: a simple and efficient method for gene transfer. *Mol. Genet. Genomics* **271**, 50–59 (2004).
58. Shen, J. R., Kawakami, K. & Koike, H. Purification and crystallization of oxygen-evolving photosystem II core complex from thermophilic cyanobacteria. *Methods Mol. Biol.* **684**, 41–51 (2011).
59. Sugiura, M. & Inoue, Y. Highly purified thermo-stable oxygen-evolving photosystem II core complex from the thermophilic cyanobacterium *Synechococcus elongatus* having His-tagged CP43. *Plant Cell Physiol.* **40**, 1219–1231 (1999).
60. 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.* **29**, 1233–1239 (1988).
61. 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).
62. Schagger, H. Tricine-SDS–PAGE. *Nat. Protoc.* **1**, 16–22 (2006).
63. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
64. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
65. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
66. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
67. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
68. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
69. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

## Acknowledgements

We thank J. Lei and the staff at the Tsinghua University Branch of the National Center for Protein Sciences Beijing for providing facility support, the Explorer 100 cluster system of the Tsinghua National Laboratory for Information Science and Technology for providing computation resources, and H. Deng and X. Meng in the Proteinomics Facility at the Technology Center for Protein Sciences, Tsinghua University, for protein MS analysis. This work was supported by the National Key R&D Program of China (grant nos. 2017YFA0503700, 2016YFA0501101, 2017YFA0504600, 2020YFA0907600 and 2019YFA0906300), the National Natural Science Foundation of China (grant no. 31470339), the Strategic Priority Research Program of CAS (grant nos. XDA27050402 and XDB17000000), a CAS Key Research programme for Frontier Science (grant no. QYZDY-SSW-SMC003), Youth Innovation Promotion Association of CAS (grant no. 2020081) and CAS Interdisciplinary Innovation Team (grant no. JCTD-2020-06).

## Author contributions

G. Han, J.-R.S. and S.-F.S. conceived the project. Y.X., G. Han and Q.Z. performed the sample isolation and characterization. G. Huang took the cryo-EM images, processed the cryo-EM data and built the structure model. Y.X., G. Huang, X.Y., W.W., G. Han, S.-F.S. and J.-R.S. analysed the structure. Y.X., G. Huang, G. Han, J.-R.S. and S.-F.S. jointly wrote the manuscript, and all the authors contributed to the discussions of the results.

## Competing interests

The authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41477-021-00961-7>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41477-021-00961-7>.

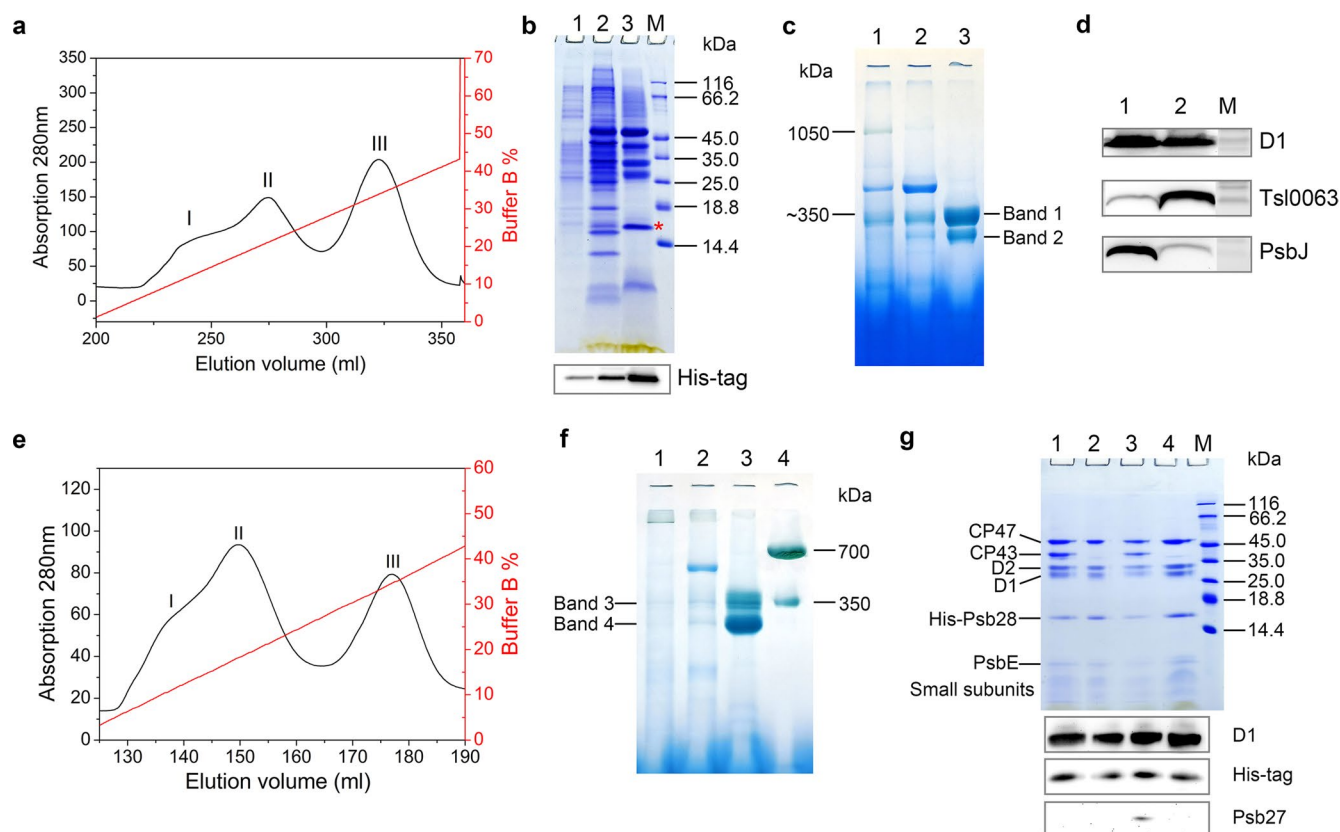
**Correspondence and requests for materials** should be addressed to G.H., S.-F.S. or J.-R.S.

**Peer review information** *Nature Plants* thanks Rob Burnap, Nathan Nelson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

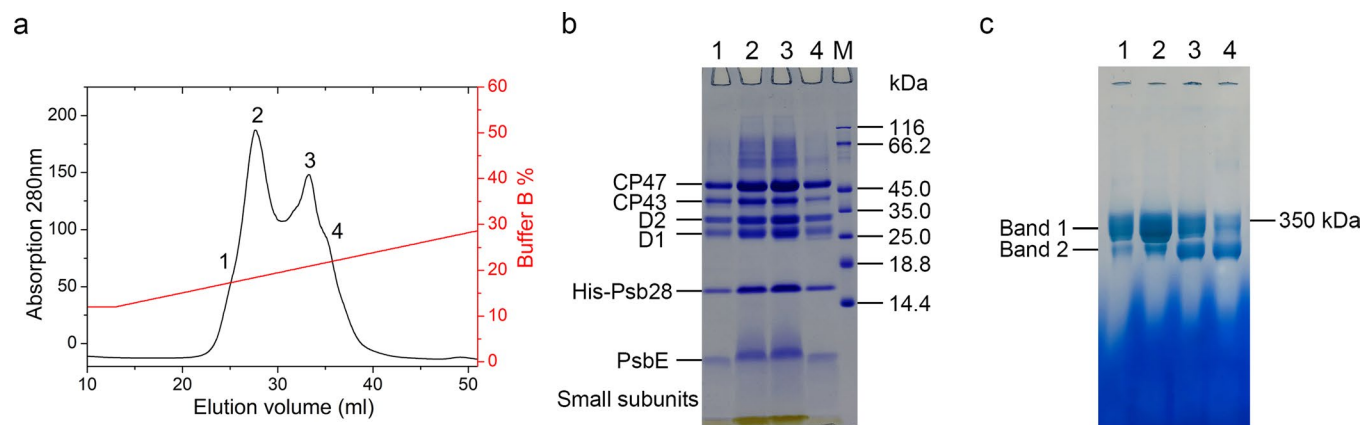
**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021



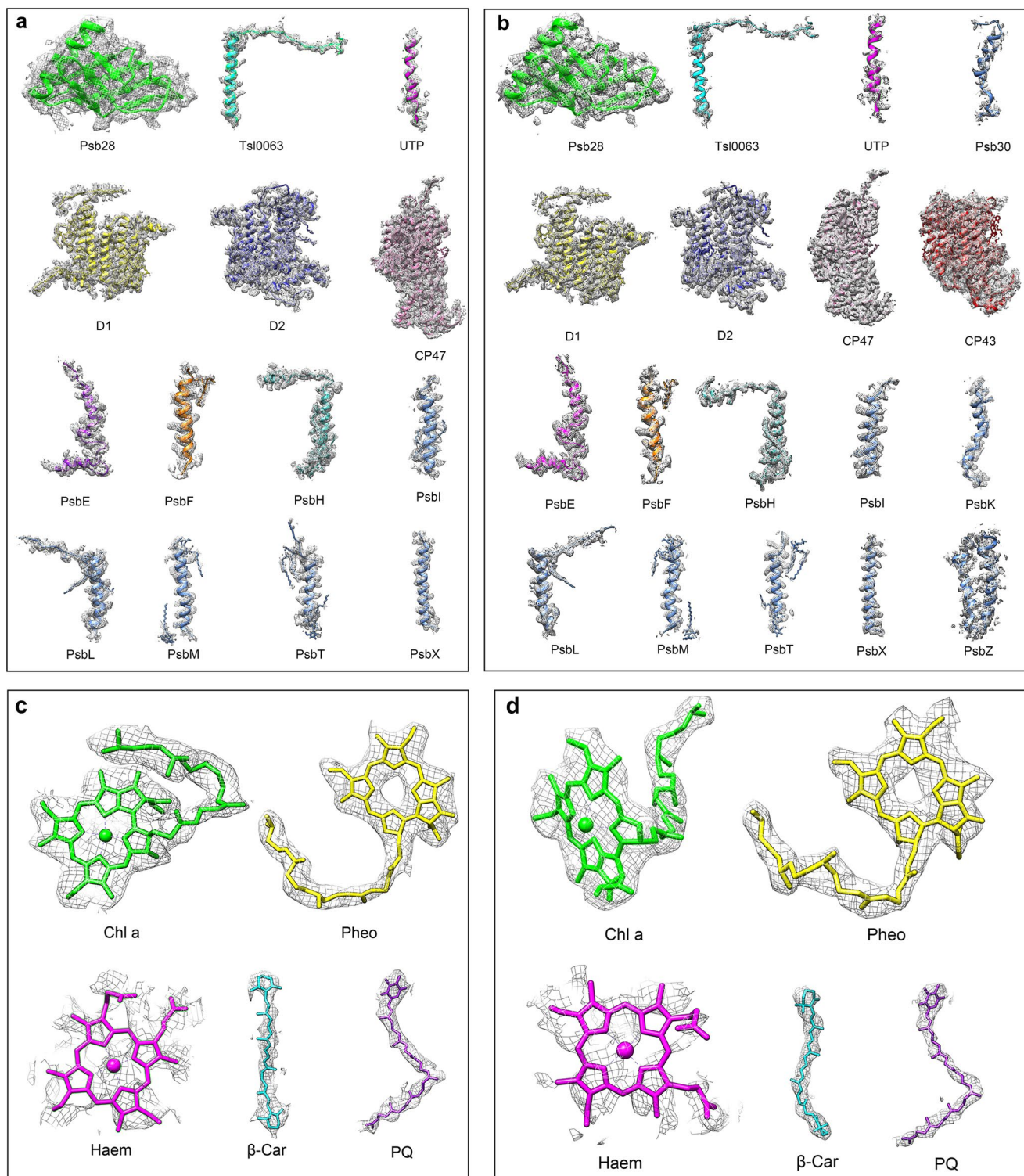
**Extended Data Fig. 1 | Purification and characterization of Psb28-PSII and Psb28-RC47 complexes from the HisPsb28- $\Delta$ PsbV and HisPsb28-WT strain of *T. vulcanus*.** **a** Elution pattern of the Psb28-PSII and Psb28-RC47 complexes from the HisPsb28- $\Delta$ PsbV strain by Ni affinity chromatography column. **b** SDS-PAGE analysis of the purified samples shown in panel **a**. Lane 1: fraction I; lane 2: fraction II; lane 3: fraction III; lane M: Molecular weight marker. The band of Psb28 was indicated by a red star. Western-blotting analysis of His-Psb28 was shown in the bottom of the SDS-PAGE. **c** BN-PAGE analysis of the purified samples after the Ni affinity chromatography in panel **a**. Lane 1: fraction I; lane 2: fraction II; lane 3: fraction III. **d** Western-blotting analysis of native PSII and fraction III of panel **a** using the antibody against D1, Tsl0063 and PsbJ subunits. Lane 1: native PSII; lane 2: fraction III of panel **a**; lane M: Molecular weight marker. **e** Elution pattern of the Psb28-PSII and Psb28-RC47 complexes from the HisPsb28-WT strain by Ni affinity column. **f** BN-PAGE analysis of the purified samples shown in panel **e**. Lane 1: fraction I; lane 2: fraction II; lane 3: fraction III; lane 4: native PSII dimer with some contaminations of the PSII monomer. **g** Two-dimensional SDS-PAGE analysis of band 1 to band 4 shown in panel **c** and **f**. Lanes 1-2: band 1-2 from lane 3 of panel **c**; lanes 3-4: band 3-4 from lane 3 of panel **f**; lane M: Molecular weight marker. Western-blotting analyses with the antibodies against D1, His-tag and Psb27 were shown in the bottom of the SDS-PAGE. The primary antibody against D1 (AS05084) was purchased from Agrisera. The primary antibody against His-tag (A02051) was purchased from Abbkine. The primary antibody against PsbJ and Tsl0063 were custom-made by Genscript, respectively. The horseradish peroxidase (HRP)-conjugated secondary antibody (Goat Anti Rabbit IgG) (AS09602) was purchased from Agrisera. Data shown in this figure is repeated more than three times, and all resulted in the same results.



**Extended Data Fig. 2 | Further purification and characterization of Psb28-PSII and Psb28-RC47 complexes from the HisPsb28- $\Delta$ PsbV strain of**

***T. vulcanus*.** **a** Elution pattern of the Psb28-PSII and Psb28-RC47 complexes (fraction III in Extended Data Fig. 1a) from a Mono Q anion-exchange column.

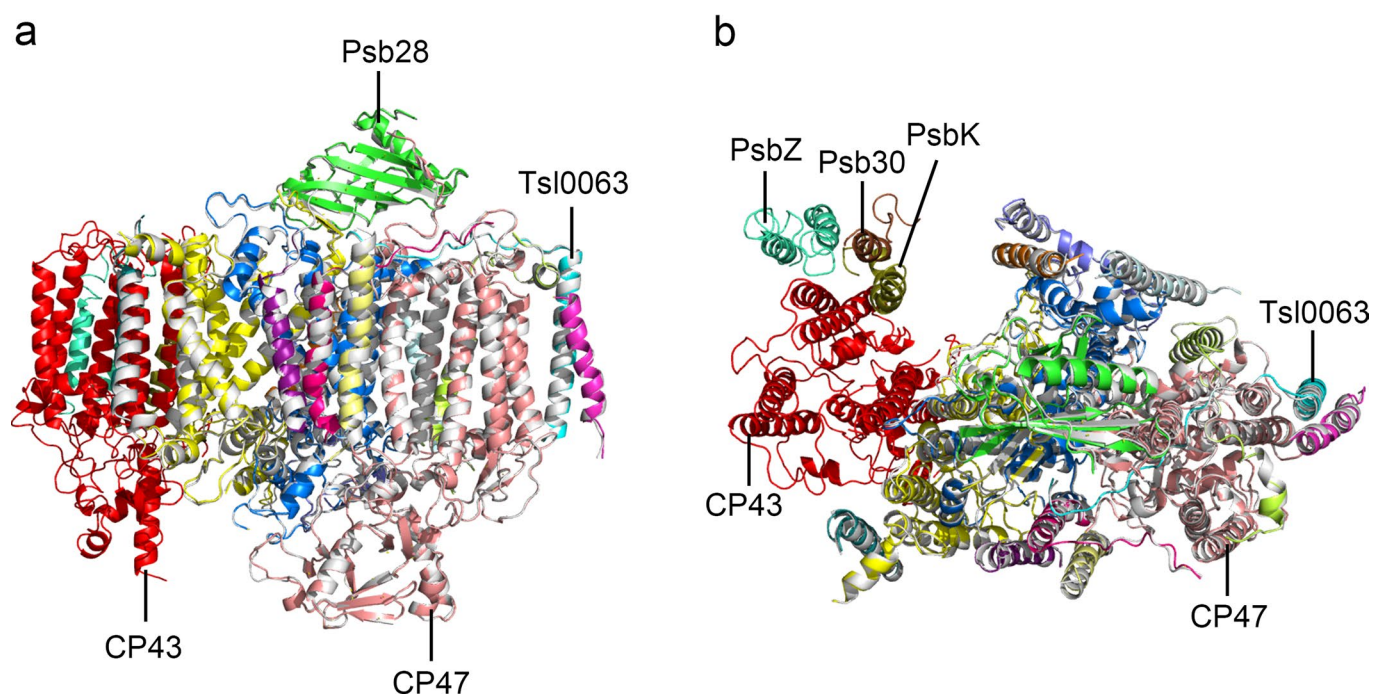
**b** SDS-PAGE analysis of the purified samples shown in panel **a**. Lane 1: fraction 1; lane 2: fraction 2; lane 3: fraction 3; lane 4: fraction 4; lane M: Molecular weight marker. **c** BN-PAGE analysis of the purified samples after the Mono Q anion-exchange column. Lane 1: fraction 1; lane 2: fraction 2; lane 3: fraction 3; lane 4: fraction 4. Data shown in this figure is conducted more than three times, and all resulted in the same results.



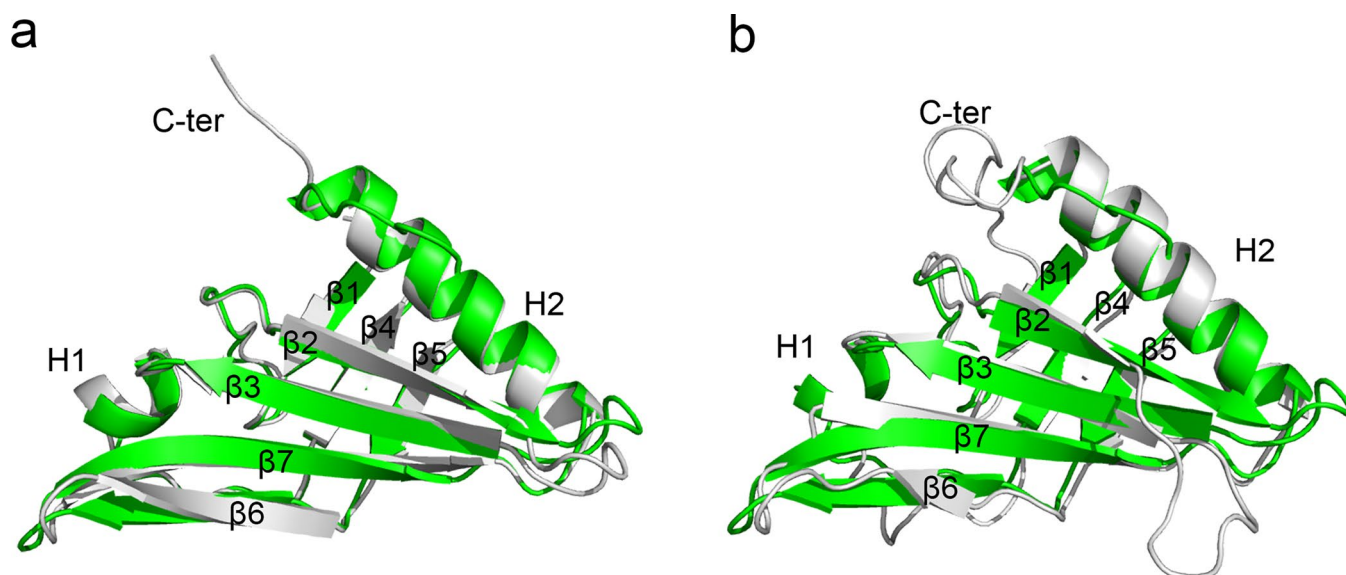
**Extended Data Fig. 3 | Cryo-EM densities and structural models of the PSII core intrinsic, extrinsic subunit and various cofactors of the Psb28-RC47 complex (a) and Psb28-PSII complex (b).** The intrinsic subunits and extrinsic subunit are shown as mixed cartoon/stick model. All cofactors are shown as stick models. Fe and Mg are shown as magenta and green spheres respectively. The cryo-EM density map of each subunit is depicted in gray meshes.

|               | 1    | 10            | 20                   | 30 | 40         | 50  |
|---------------|------|---------------|----------------------|----|------------|-----|
| tr V5V6Y3     | MRYT | DEGGRLNNFAIEP | KVYQAQPWTPQOKVRAALLV | G  | GGLLLVAGLV | IAI |
| tr A0A3M2FSJ2 | MRYT | DEGGRLNNFAIEP | KVYQAQPWTPQOKVRAALLV | G  | GGLLLVAGLV | IAI |
| tr Q8DMP8     | MRYT | DEGGRLNNFAIEP | KVYQAQPWTPQOKVRAALLV | G  | GGLLLVAGLV | IAI |
| tr A0A5C2M567 | MRYT | DEGGRLNNFAIEP | KVYQAQPWTPQOKVRAALLV | V  | GGLLLVAGLV | IAI |
| tr A0A3B7MFR6 | MRYT | DEGGRLNNFAIEP | KVYQAQPWTPQOKVRAALLV | V  | GGLLLVAGLV | IAI |

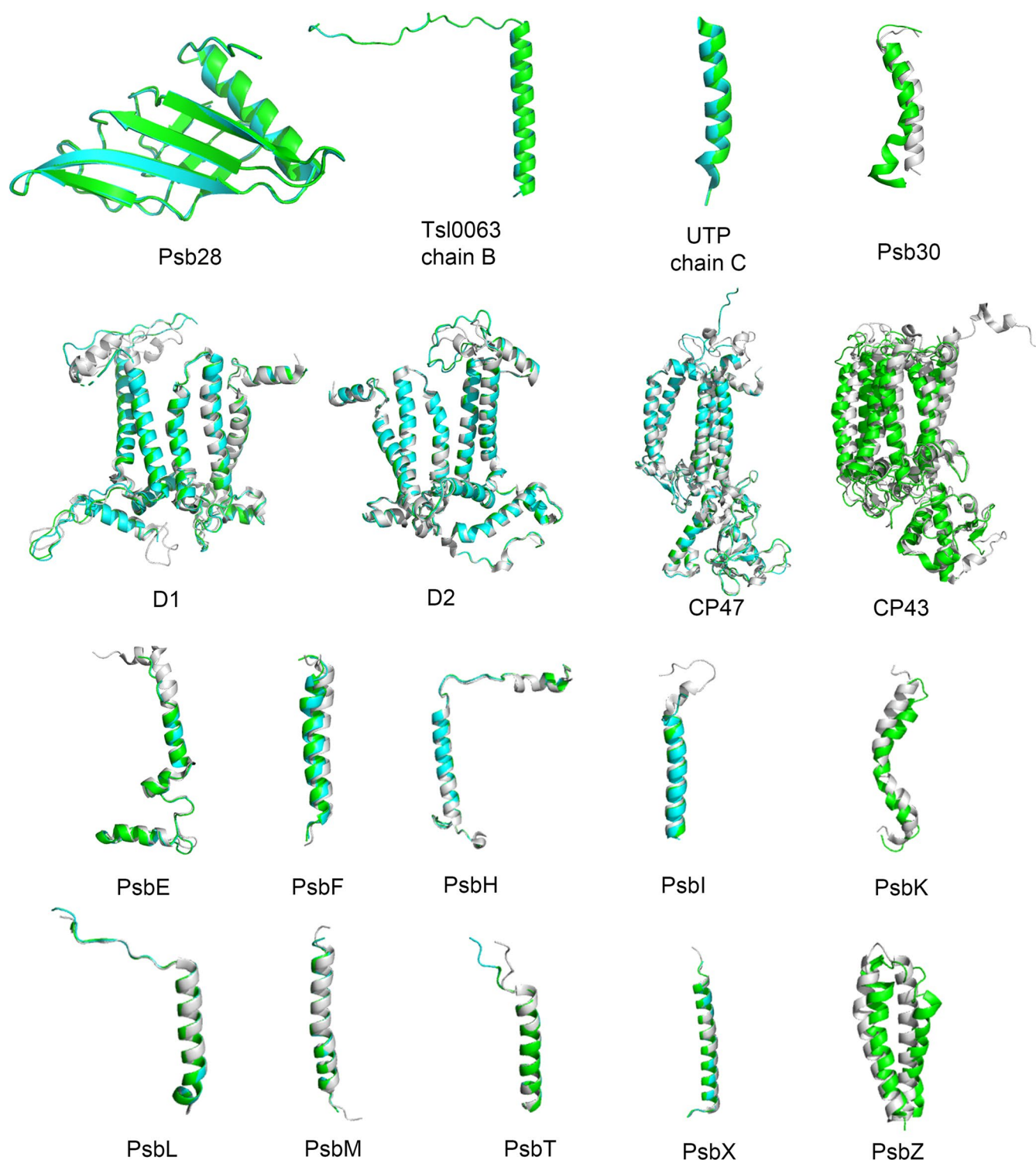
**Extended Data Fig. 4 | Sequence comparison of Tsl0063 with other subunits that have a sequence similarity above 90%.** V5V6Y3: CAB/ELIP/HLIP family protein from *Thermosynechococcus* sp. NK55a; A0A3M2FSJ2: light-harvesting-like protein (Ssl1498 family) from *Cyanobacteria bacterium* J003; Q8DMP8: Tsl0063 protein from *Thermosynechococcus elongatus* strain BP-1; A0A5C2M567: light-harvesting-like protein (Ssl1498 family) from *Thermosynechococcus* sp. CL-1; A0A3B7MFR6: light-harvesting-like protein (Ssl1498 family) from *Thermosynechococcus elongatus* PKUAC-SCTE542. (These information is obtained by searching the database of <https://www.uniprot.org/>).



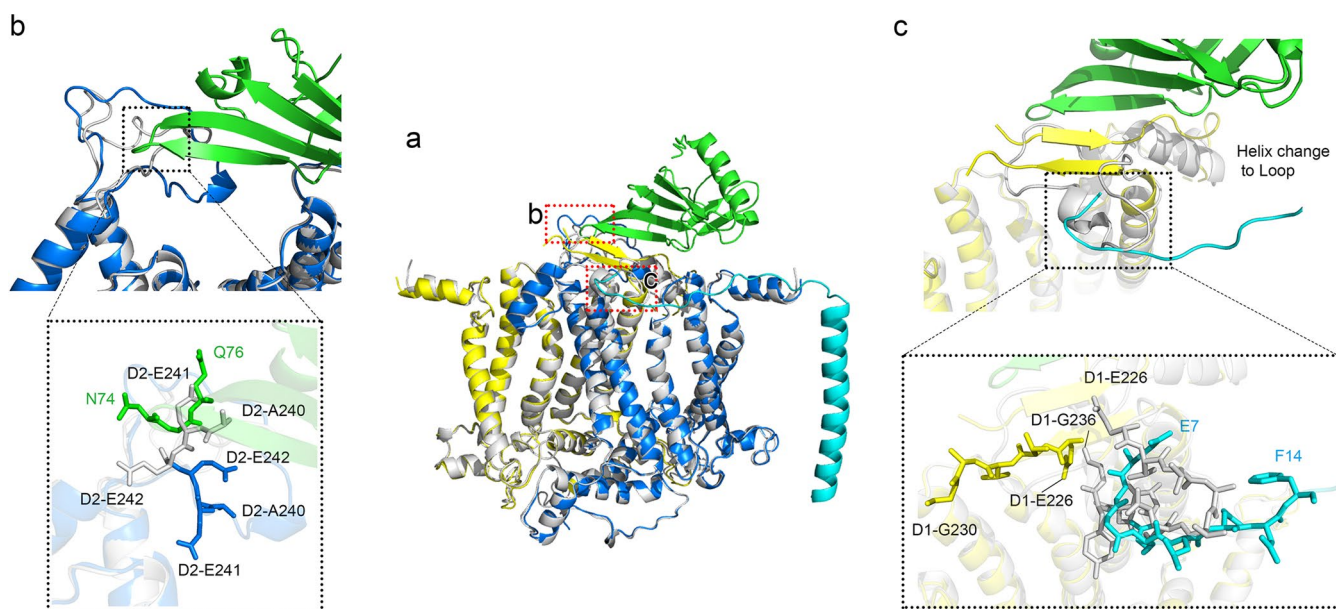
**Extended Data Fig. 5 | Structural comparison of the Psb28-RC47 and Psb28-PSII complexes of *T. vulcanus*.** **a, b** Structural comparison of the Psb28-RC47 and Psb28-PSII complexes of *T. vulcanus*, viewed along the membrane plane (**a**) and from the stromal side (**b**). All subunits of Psb28-RC47 are shown in grey and the subunits of Psb28-PSII are colored differently as those in Fig. 1.



**Extended Data Fig. 6 | Structural comparison of the Psb28 subunit from the Psb28-RC47 complex of *T. vulcanus* with its crystal structure and NMR structure. a, b** Structural comparison of the Psb28 subunit from the Psb28-RC47 complex of *T. vulcanus* with its crystal structure from *T. elongates* (PDB ID code: 3ZPN) (**a**) and NMR structure from *Synechocystis* sp. strain PCC 6803 (PDB ID code: 2KVO) (**b**). Psb28 from Psb28-RC47 is shown in green, and its crystal structure or NMR structure is shown in grey.



**Extended Data Fig. 7 | Structural comparison of the subunits between Psb28-RC47, Psb28-PSII and native PSII (PDB: 3WU2).** Subunits in Psb28-RC47, Psb28-PSII and native PSII are shown in cyan, green and grey, respectively.



**Extended Data Fig. 8 | Structural conflicts between D2 and Psb28, D1 and Tsl0063.** **a** Overall structure of D1, D2, Psb28 and Tsl0063 in Psb28-PSII and native PSII (PDB: 3WU2). **b** Structural conflicts between D2 and Psb28. **c** N-terminal region (E7-F14) of Tsl0063 (Psb28-PSII) and the D-E region (E226-G236) of D1 from native PSII. Enlarged view of the boxed area shows the amino acid residues involved in the structural conflicts. D1 subunits in Psb28-PSII and native PSII are shown in yellow and grey, respectively; D2 subunits in Psb28-PSII and native PSII are shown in marine and grey, respectively; Psb28 and Tsl0063 subunits in Psb28-PSII is shown in green and cyan, respectively.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☒ ☐ The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- ☒ ☐ A description of all covariates tested
- ☒ ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☒ ☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☒ ☐ For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection serialEM 3.8

Data analysis Proteome Discovery searching algorithm 1.4, SerialEM 3.8, MotionCor2 1.1.0, CTFFIND 4.0, RELION-3.0, Chimera 1.14, Coot 0.8.9, Phenix 1.18, PyMOL 2.1.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The cryo-EM density map and atomic models have been deposited in the Electron Microscopy Data Bank and the Protein Data Bank (EMD ID code 30902 and PDB ID code 7DXA for Psb28-RC47, EMD code 30909 and PDB code 7DXH for Psb28-PSII). The data that support the findings of this study are available from the corresponding authors upon reasonable request.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

|                 |                                                                                                                                                                                                                                                                                                                                                                                                               |
|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sample size     | No statistic method was used to predetermine the sample size. The amount of cryo-EM micrographs collected was based on the previous knowledge and the results indicated that the size is sufficient to generate a density map at 3.14 Å resolution.                                                                                                                                                           |
| Data exclusions | No data exclusion criteria was pre-established. 2D and 3D classifications were used to obtain multiple classes. Only the particles in the classes that showed clear structural signals and intact structures were chosen, combined and used in the final reconstruction of cryo-EM map and refinement. Details are described in the flowchart of Supplementary Figs. 5, 6, Supplementary Table 2 and Methods. |
| Replication     | Multiple rounds of structural refinement have been performed and all resulted in same density maps. Most experiments concerning the purification and characterization of Psb28-RC47 and Psb28-PSII have been repeated independently for more than three times, and all reproduced the same results successfully.                                                                                              |
| Randomization   | Randomization is not relevant to our study because we aimed to solve the structure of the specific protein supercomplexes.                                                                                                                                                                                                                                                                                    |
| Blinding        | Blinding is not relevant to our study because we aimed to solve the structure of the specific protein supercomplexes.                                                                                                                                                                                                                                                                                         |

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a                                 | Involved in the study                                  |
|-------------------------------------|--------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |

### Methods

| n/a                                 | Involved in the study                           |
|-------------------------------------|-------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Antibodies

|                 |                                                                                             |
|-----------------|---------------------------------------------------------------------------------------------|
| Antibodies used | The primary antibody were either purchased from Agrisera or custom-made.                    |
| Validation      | The antibodies were validated with native PSII samples by immunoblotting for several times. |