Published in final edited form as:

Nat Plants. 2021 April ; 7(4): 524–538. doi:10.1038/s41477-021-00895-0.

Structural insights into photosystem II assembly

Jure Zabret[1,14], **Stefan Bohn**[2,14], **Sandra K. Schuller**[3,4], **Oliver Arnolds**[5], **Madeline Möller**[1], **Jakob Meier-Credo**[6], **Pasqual Liauw**[1], **Aaron Chan**[7], **Emad Tajkhorshid**[7], **Julian D. Langer**[6,8], **Raphael Stoll**[5], **Anja Krieger-Liszkay**[9], **Benjamin D. Engel**[2,10,11], **Till Rudack**[12,13,✉], **Jan M. Schuller**[3,4,✉], **Marc M. Nowaczyk**[1,✉]

[1]Department of Plant Biochemistry, Faculty of Biology and Biotechnology, Ruhr University Bochum, Bochum, Germany.

[2]Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, Martinsried, Germany.

[3]Department of Structural Cell Biology, Max Planck Institute of Biochemistry, Martinsried, Germany.

[4]CryoEM of Molecular Machines, SYNMIKRO Research Center and Department of Chemistry, Philipps University of Marburg, Marburg, Germany.

[5]Biomolecular Spectroscopy and RUBiospek|NMR, Faculty of Chemistry and Biochemistry, Ruhr University Bochum, Bochum, Germany.

[6]Proteomics, Max Planck Institute of Biophysics, Frankfurt, Germany.

[7]NIH Center for Macromolecular Modeling and Bioinformatics, Beckman Institute for Advanced Science and Technology, Department of Biochemistry, and Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[8]Proteomics, Max Planck Institute for Brain Research, Frankfurt, Germany.

[9]Institute for Integrative Biology of the Cell (I2BC), Université Paris-Saclay, CEA, CNRS, Gif-sur-Yvette, France.

[10]Helmholtz Pioneer Campus, Helmholtz Zentrum München, Neuherberg, Germany.

✉**Correspondence and requests for materials** should be addressed to T.R., J.M.S. or M.M.N. till.rudack@rub.de; jan.schuller@synmikro.uni-marburg.de; marc.m.nowaczyk@rub.de.

Author contributions

B.D.E., T.R., J.M.S. and M.M.N. conceived the research, prepared the figures and wrote the manuscript with the contribution of J.Z. and all other authors. M.M.N. coordinated the activities. Preparation of mutants, PSII isolation and biochemical analysis were performed by J.Z., M.M, P.L. and M.M.N. Mass spectrometry analysis was done by J.M.-C. and J.D.L. J.M.S., S.B. and B.D.E. performed the cryo-EM analysis. T.R. built the structural model with the help of S.K.S., A.C. and E.T. Fluorescence spectroscopy was carried out by J.Z. and M.M.N. EPR experiments were conducted by A.K.-L. NMR experiments were conducted and analysed by O.A. and R.S. All authors approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41477-021-00895-0.

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

[11]Department of Chemistry, Technical University of Munich, Garching, Germany.

[12]Biospectroscopy, Center for Protein Diagnostics (ProDi), Ruhr University Bochum, Bochum, Germany.

[13]Department of Biophysics, Faculty of Biology and Biotechnology, Ruhr University Bochum, Bochum, Germany.

[14]These authors contributed equally: Jure Zabret, Stefan Bohn.

Abstract

Biogenesis of photosystem II (PSII), nature's water-splitting catalyst, is assisted by auxiliary proteins that form transient complexes with PSII components to facilitate stepwise assembly events. Using cryo-electron microscopy, we solved the structure of such a PSII assembly intermediate from *Thermosynechococcus elongatus* at 2.94 Å resolution. It contains three assembly factors (Psb27, Psb28 and Psb34) and provides detailed insights into their molecular function. Binding of Psb28 induces large conformational changes at the PSII acceptor side, which distort the binding pocket of the mobile quinone (Q_B) and replace the bicarbonate ligand of non-haem iron with glutamate, a structural motif found in reaction centres of non-oxygenic photosynthetic bacteria. These results reveal mechanisms that protect PSII from damage during biogenesis until water splitting is activated. Our structure further demonstrates how the PSII active site is prepared for the incorporation of the Mn_4CaO_5 cluster, which performs the unique water-splitting reaction.

Photosystem II (PSII) is the only enzyme that catalyses the light-driven oxidation of water, a thermodynamically demanding reaction that drives photosynthesis, sustaining life on our planet[1–3]. This multisubunit membrane protein complex is located in the thylakoid membranes of cyanobacteria, algae and plants. PSII strips electrons from water and injects them into the photosynthetic electron transport chain. It forms a homodimer with a molecular mass of ~500 kDa (ref. [4]), with each monomer composed of at least 20 protein subunits and numerous cofactors, including chlorophylls, quinones, carotenoids, lipids, bicarbonate and the unique Mn_4CaO_5 cluster[5–7]. The two core proteins, D1 and D2, form a central, membrane-intrinsic heterodimer, which binds all important redox cofactors involved in internal electron transfer[8]. Light excitation leads to a charge-separated state in which an electron is transferred from the reaction centre chlorophylls P_{680} (ref. [9]) to the nearby pheophytin[10]. Subsequently, the electron is passed to the bound plastoquinone (Q_A) and then to the mobile plastoquinone molecule (Q_B), which leaves the complex after accepting two electrons and two protons[11]. The electron hole at P_{680} is filled by oxidation of an adjacent tyrosine residue (Tyr_Z)[12] and finally by the oxygen-evolving complex (OEC), which contains the Mn_4CaO_5 cluster. In cyanobacteria, the cluster is shielded on the luminal side by the three extrinsic proteins, PsbO, PsbU and PsbV, which regulate access to the OEC by forming an intricate network of channels for different substrates and products[13]. Light energy is collected and funnelled towards P_{680} by the two membrane-intrinsic antenna proteins CP43 and CP47. These proteins bind most of the chlorophyll molecules and are located on opposite sides of the D1/D2 heterodimer[14]. Moreover, at least 12 small transmembrane subunits with one or two transmembrane helices have been identified in PSII[15], including cytochrome b_{559} (ref. [16]).

Structural and spectroscopic investigations have revealed these aforementioned comprehensive insights into PSII function[17–21] but we are still far from understanding PSII biogenesis in molecular detail. How nature facilitates the assembly of a multisubunit, multicofactor membrane protein complex is a fundamental, unsolved question. The biogenesis of PSII is even more challenging, as the mature complex performs sophisticated and extreme redox chemistry to catalyse the light-driven oxidation of water. This can easily lead to the formation of reactive oxygen species (for example, singlet oxygen is produced by triplet chlorophyll in the PSII reaction centre) and subsequent loss of function due to damaged proteins and cofactors[22,23]. Biogenesis intermediates with only partially functional fragments of the redox chain are particularly prone to damage, thus demanding specialized protection mechanisms for the assembly process[24]. Therefore, PSII biogenesis is not a spontaneous process but rather must be tightly regulated by the action of assembly factors. Thus far, >20 auxiliary proteins have been identified that guide the stepwise assembly of PSII subunits and cofactors via intermediate modules, which are assembled independently and then joined together to produce mature PSII[25–27]. In cyanobacteria, PSII biogenesis begins with the formation of the D1/D2 heterodimer reaction centre (RC) complex from the D1 precursor protein (pD1) and the D2 protein. This is assisted by the PSII assembly factor Ycf48 after partial processing of the D1 C-terminal extension by the D1-specific peptidase CtpA[28,29]. In the next step, the assembly factor Psb28 helps CP47 join the RC complex to form the RC47 complex, in which the intermediate iD1 is further processed into its mature form by CtpA[30,31]. Almost all ligands of the Mn_4CaO_5 cluster are already present at this stage, except for those provided by CP43, which comes preconstructed with assembly factor Psb27 and several small subunits (together called the CP43 module)[32]. Psb28 is released as CP43 binds and the resulting Psb27-PSII monomer is activated by maturation of the OEC and the binding of the extrinsic proteins PsbO, PsbU and PsbV[33–35]. OEC assembly is a multistep process that requires a functional upstream redox chain for the oxidation of Mn^{2+} to build up the cluster's μ-oxo bridges between the manganese atoms[36,37]. The mechanistic and structural details of this photoactivation process are not yet understood. In the consensus 'two quantum model'[38–40], a single Mn^{2+} ion bound to the high-affinity site (HAS)[41] is oxidized to Mn^{3+}. This initiating, light-dependent step is followed by a slow, light-independent phase and further fast, light-dependent steps in which the remaining Mn^{2+} ions are oxidized and incorporated. Understanding the light-independent slow phase is key to unravelling the mechanism of photoactivation. Finally, PSII biogenesis completes with dimerization of two fully assembled monomers and attachment of the soluble phycobilisome antenna complexes. Interestingly, deletion of *psbJ*, which encodes a small single transmembrane helix protein at the entrance of the PSII plastoquinone channel, decreases PSII-mediated oxygen evolution and increases the lifetime of the reduced primary acceptor (Q_A^-) in cyanobacterial and tobacco chloroplasts[42]. These effects are caused by a massive accumulation of an intermediate monomeric PSII complex, which contains both assembly factors Psb27 and Psb28 (ref. [43]). Physiological studies of Psb27- and Psb28-deletion strains point towards multifaceted functions. Cyanobacterial mutants lacking Psb28 exhibited slower autotrophic growth, particularly under stress conditions[31,44] and limited synthesis of chlorophyll-binding proteins but without decreased PSII functionality[31]. The Psb28 mutant also exhibited an overall increase in PSII repair and faster recovery from photodamage[31]. Chemical cross-linking combined with mass spectrometry revealed that Psb28 binds to the cytosolic side of CP47, close to cytochrome b_{559} and the Q_B binding site. On the basis of this, researchers postulated a protective role for Psb28, where it blocks electron transport to the acceptor side of PSII to shield the RC47 complex from excess photodamage during the assembly process[45]. This hypothesis is strengthened by the observation that Psb28 is also found in PSII repair complexes[46]. The luminal PSII assembly factor Psb27 has been similarly well investigated. This lipoprotein is predominantly associated with inactive PSII fractions involved in assembly or repair[32,34,46–50], stabilizing the CP43 luminal domain and presumably facilitating the assembly of the OEC.

Our current knowledge of PSII biogenesis mainly describes the order of events and protein composition of each intermediate, as well as the general roles of PSII assembly factors. However, the precise molecular functions of these intermediate complexes and the involved assembly factors are still elusive due to their low abundance and intrinsic instability. High-resolution structural information is of vital importance to gain a deeper understanding of the molecular action of PSII assembly factors, as they are proposed to alter the structures of their associated PSII proteins to provide protection or facilitate specific biogenesis transitions.

Here, we use cryo-electron microscopy (cryo-EM) single-particle analysis to describe the molecular structure of a PSII assembly intermediate. This structure represents one of the key transitions in PSII biogenesis: the attachment of the CP43 module to the pre-assembled RC47 complex, which precedes incorporation and activation of the Mn_4CaO_5 cluster. We complement this structural data with spectroscopic analysis, revealing detailed insights into the molecular mechanisms of PSII assembly. Our study provides mechanistic answers to three long-standing questions: (1) How do assembly factors modulate the structures of PSII subunits to assist biogenesis? (2) How is PSII protected from photodamage during assembly? (3) How is the PSII active site prepared for incorporation of the Mn_4CaO_5 cluster?

Results

Stable PSII intermediates were purified from the *Thermosynechococcus elongatus* Δ*psbJ* mutant[43] by affinity chromatography using a Twin-Strep-tag fused to the C terminus of the CP43 subunit and subsequent ion exchange chromatography (IEC) (Supplementary Fig. 1a). The main peak of the IEC profile corresponds primarily to monomeric PSII, which lacks the extrinsic subunits PsbO, PsbU and PsbV that are indicative for water-splitting activity (Supplementary Fig. 1b,c). Single-particle cryo-EM analysis of this PSII fraction resulted in three different high-resolution maps that allowed model building with high confidence and excellent statistics (Supplementary Fig. 2 and Supplementary Table 1). In addition to the protein subunits, we also assigned all essential non-protein cofactors, including chlorophylls, quinones, carotenoids and lipids, which are also present in the mature PSII complex (Supplementary Fig. 3). Consistent with previous biochemical studies[34,43,48], the EM density corresponding to the fully assembled, active Mn_4CaO_5 cluster is missing in the purified biogenesis intermediates. The first cryo-EM map (2.94 Å), which we call PSII-I (for PSII intermediate), provides a snapshot of the attachment of the CP43 module to the pre-assembled RC47 complex (Fig. 1). This PSII-I contains three assembly factors (Psb27, Psb28 and Psb34), as well as almost all the membrane-intrinsic subunits and cofactors found in mature PSII, although CP43 is bound in an immature conformation. Psb27 and Psb28 are well-known assembly factors[31,32,35,43,51], whereas the additional single transmembrane helix protein (tsl0063), which we named Psb34, has not been described before. The small subunit PsbY, which is known to be loosely bound[52], is not resolved in our structure. In addition, PsbJ is not present, as the corresponding gene was inactivated to stall PSII assembly at this specific transition[43].

The two additional maps serve as internal controls. PSII-I′ (2.76 Å) lacks Psb27 but is otherwise comparable to PSII-I; the root mean square deviation of the C_α atomic positions between similar subunits of the two complexes is 0.4 Å. Probably, Psb27 was partly lost during sample preparation. The third cryo-EM map (2.82 Å), which we call PSII-M (for PSII monomer), represents a monomeric PSII complex without bound assembly factors. Comparison of our PSII-M structure with a crystal structure of monomeric PSII[52] (PDB-ID 3KZI, 3.6 Å) reveals only minimal differences between both structures, with a C_α root mean square deviation of 1.3 Å, which verifies that the structural changes observed in PSII-I are not caused by the deletion of PsbJ.

Psb34 binds to CP47 and is involved in the RC47 transition.

Our PSII-I structure provides identification of the single transmembrane helix protein Psb34 bound to a PSII assembly intermediate (Fig. 2a), which we also confirmed by mass spectrometry (Fig. 2b). Psb34 was probably overlooked previously due to its hydrophobicity and small size. It has a single transmembrane helix that binds to the CP47 antenna protein in close proximity to PsbH (Fig. 2a). Its conserved long N-terminal arm is located at the side and top of the D2 subunit (Fig. 2a). In addition, we independently confirmed the interaction of Psb34 with PSII assembly intermediates by isolation of Strep-tagged Psb34 complexes from otherwise unmodified cells, indicating a specific function of Psb34 in the attachment of CP43 to RC47 (Fig. 2c). In this case, two distinct PSII intermediates were isolated via pulldown of Strep-tagged Psb34: the RC47 complex with bound Psb28 and the subsequent PSII-I after attachment of CP43 and Psb27 (Fig. 2c). This observation implies that Psb28 is usually released from the PSII-I after attachment of CP43, probably after incorporation of PsbJ, as this trigger is missing in the analysed Δ*psbJ* mutant. PSII-I, which only accumulates in the Δ*psbJ* mutant, seems to represent a hybrid state of both PSII intermediates that might be stabilized by the incomplete binding of CP43. Psb34 shows sequence similarity to high-light inducible proteins (HLIPs), which play a role in transient chlorophyll storage and chlorophyll biosynthesis[53]. However, the chlorophyll-binding motive is missing in Psb34 (Supplementary Table 2), suggesting a distinct function for this protein in PSII biogenesis.

Psb28 induces conformational changes in the D1 D-E loop.

Psb28 binds on the cytosolic faces of the D1 and D2 subunits, directly above the Q_B binding site (Fig. 3a), which differs from the position that was previously predicted by mass spectrometry[45]. Its binding induces the formation of an extended β-hairpin structure that incorporates the central antiparallel β-sheet of Psb28, the C terminus of CP47 and the D1 D-E loop[54] (Fig. 3a). Binding of Psb28 to the C terminus of CP47 also imparts a directionality to the assembly process. In the Psb28-free complex (PSII-M), the CP47 C terminus blocks the Psb28 binding site by interacting with the D1 D-E loop, thus preventing the reverse process and perturbation of active PSII by Psb28. Using nuclear magnetic resonance (NMR) spectroscopy, we performed chemical shift perturbation (CSP) experiments with recombinant Psb28 and a synthetic peptide of the conserved CP47 C terminus to characterize this interaction in detail and determine the dissociation constant (K_D) (Fig. 3 and Supplementary Fig. 4). The CSP measurements indicated notable shifts with a chemical shift difference ($\Delta\delta$) of >1 s.d. located at strands β3 and β4 as well at the C-terminal region of Psb28 (Fig. 3c,d). On peptide binding, resonances for several residues gradually appeared with increasing peptide concentration, which were line-broadened beyond detection for the free form of Psb28. This observation indicates a less dynamic and more rigid complex structure. This is further supported by the heteronuclear Overhauser effect (NOE) data, which show that the C terminus of Psb28 becomes rigid from L108 to K112 on CP47 peptide binding due to creation of an intermolecular β-sheet (Fig. 3e). Two-dimensional lineshape analysis was performed, yielding a K_D of 53.92 ± 0.41 μM and a dissociation rate k_{off} of $10.14 \pm 0.16\ s^{-1}$, which is consistent with the observed slow exchange in the NMR spectra (Fig. 3b). The affinity of Psb28 for full-length CP47 and PSII might indeed be even higher due to additional contacts between Psb28 and the D-E loop of D1 (Fig. 3a).

Psb28 prevents full association of CP43 and distorts the QB-site.

Binding of Psb28—with support of Psb34—causes major structural perturbations at the PSII acceptor side (Supplementary Videos 1 and 2), which mainly involve the D-E loops of the central D1 and D2 subunits. Comparison of the CP43 structure in PSII-I with that in our PsbJ-free control PSII-M (Fig. 4a–d) or with that in mature monomeric PSII (PDB-ID 3KZI) (Fig. 4c,d) reveals several differences. The CP43 C terminus is not resolved in PSII-I, probably due to an immature position of the last transmembrane helix of CP43 and an altered conformation of the D1 D-E loop, which may prevent binding of the CP43 C terminus to the cytoplasmic PSII surface (Fig. 4b). This region is close to the loop between helices D and E of the D2 subunit, which is also altered by binding of Psb28, as clearly shown by movement of D2 Arg233 (Fig. 4b and Supplementary Fig. 5a,b). After dissociation of Psb28, the CP43 module undergoes a rigid body rotation where it clicks into place (Fig. 4b–d and Supplementary Video 1), whereas binding of PsbJ and the extrinsic proteins PsbO, PsbV and PsbU during further maturation has very little influence on the CP43-binding position (Fig. 4c,d). The part of PSII that originates from RC47 shows almost no difference between PSII-I and mature PSII (Fig. 4d), except for PsbE, which binds adjacent to PsbJ (Fig. 4c).

Most importantly, the structural changes in the D1 D-E loop may have a direct functional impact on PSII electron transfer (Fig. 4e,f), as this region coordinates several important PSII cofactors. In functional PSII, after charge separation at the RC P_{680}, electrons are transferred via pheophytin to the bound plastoquinone (Q_A) and further to mobile plastoquinone (Q_B). In our PSII-I structure, the Q_A site is fully assembled and a well-resolved Q_A molecule is bound (Fig. 4e,f and Supplementary Fig. 5c,d). The nearby non-haem iron is also already in place in PSII-I (Fig. 4e,f and Supplementary Fig. 5e,f). The Q_B binding site of the PSII-M control is comparable to mature PSII, although it is not occupied by Q_B in our preparation (Supplementary Fig. 5g). In contrast, the Q_B binding site of PSII-I is immature due to the Psb28- and Psb34-induced structural changes in the D1 D-E loop (Fig. 4e,f and Supplementary Fig. 5h). Notably, D1 Phe265, which coordinates the head group of Q_B in mature PSII, is clearly at a different position[21] (Supplementary Video 2).

Binding of Psb28 protects PSII during biogenesis.

A more detailed analysis of the structural environment close to the Q_A/Q_B binding sites revealed differences in the coordination and the hydrogen-bond network of the adjacent non-haem iron, which also indicate functional consequences for PSII electron transfer and charge recombination processes. In mature PSII, the non-haem iron is coordinated by four histidine residues and bicarbonate as the fifth ligand (Fig. 5a,c), whereas in PSII-I, the bicarbonate molecule is replaced by the E241 side chain of D2 (Fig. 5b,d, Supplementary Fig. 5e,f and Supplementary Video 3). Other residues, including D1 E244 and Y246, which bind to the bicarbonate molecule in mature PSII (Fig. 5a), are also displaced in PSII-I due to the conformational change of the D1 D-E loop (Fig. 5b, Supplementary Fig. 5e,f and Supplementary Video 3). Binding of bicarbonate is important for PSII efficiency[55], as it lowers the redox potential of (Q_A/Q_A^-) to favour forward electron transport[56,57]. If charge recombination occurs, the lower redox potential favours indirect charge recombination via $P_{680}^{\bullet+}/Pheo^{\bullet-}$. This back reaction yields triplet chlorophyll and subsequently singlet oxygen[57], a highly reactive oxygen species. Changes in the redox potential of (Q_A/Q_A^-) have been proposed to tune the efficiency of PSII depending on the availability of CO_2 as the final electron acceptor and thereby protect PSII under low CO_2 conditions[57]. Therefore, we used flash-induced variable fluorescence to monitor the stability of Q_A^- in active PSII, in the PSII-I assembly intermediate and in inactivated, Mn-depleted PSII, both of which lack a functional OEC (Fig. 5e and Supplementary Fig. 6a,b). The results indicate that the unique fluorescence traces observed with PSII-I are a consequence of Q_A/Q_A^- redox tuning by an absent OEC, an immature Q_B binding pocket and probably the replacement of bicarbonate by glutamate as a ligand of the non-haem iron (Supplementary Discussion). To determine the functional consequences of this redox tuning in PSII-I, we measured the formation of $^1O_2$ by electron paramagnetic resonance (EPR) spectroscopy using the spin probe TEMPD. The data clearly show that $^1O_2$ formation is reduced by ~30% in PSII-I compared to inactivated PSII (Fig. 5f).

Psb27 binds to loop E of CP43 at the luminal PSII surface.

Psb27 binds to the luminal side of the PSII complex, adjacent to loop E of the CP43 subunit (Fig. 6a,b). We could successfully model the Psb27 N terminus until Cys22, which is the first residue after N-terminal processing of Psb27 (ref. [34]). Psb27-Cys22 is located close to the membrane plane but the lipid moiety is not well enough resolved to build a model. In contrast to previously proposed models[58,59], the binding site of Psb27 has little overlap with the binding sites of the extrinsic subunits (PsbO, PsbV and PsbU) and has at least no direct contact to the Mn_4CaO_5 cluster binding site, particularly not to the D1 C terminus (Fig. 6a,b), which has been suggested previously[58,60]. Instead, Psb27 is bound at a remote position that might be occupied by CyanoQ in the mature complex[61]. This position of Psb27 may still prevent proper binding of PsbO[34] by lowering its binding affinity due to the partial overlap of the binding sites, although both proteins can bind together, as PsbO-containing Psb27-PSII complexes have been isolated via His-tagged Psb27 (ref. [49]). Additionally, Psb27 might stabilize loop E of CP43 in the unassembled state and facilitate its binding to the D1 subunit for the assembly of the Mn_4CaO_5 cluster[62]. This is of particular importance, as loop E of CP43 provides Arg345 and Glu342 (Arg357 and Glu354 in previous publications, see caption of Fig. 6 for details) to position the Mn_4CaO_5 cluster in mature PSII (Fig. 6b, dashed box). Glu342 serves as a ligand, whereas Arg345 stabilizes the Mn_4CaO_5 cluster by coordinating Asp170 of D1 through a hydrogen bond. Moreover, in the Psb27-bound state (PSII-I), the D1 C terminus, which is directly involved in coordination of the Mn_4CaO_5 cluster[21], is bound away from the cluster (Fig. 6c and Supplementary Fig. 7), which might be an additional consequence of Psb27-induced stabilization of the CP43 E loop. Thus, our PSII-I structure reveals not only how the Psb27 protein binds to CP43 and thus stabilizes it but also indicates a role for the maturation of the oxygen-evolving complex that is consistent with functional data from previous studies[32,35,48,62].

PSII-I contains a single cation at the immature OEC site.

The unique Mn_4CaO_5 cluster is a key feature of PSII that splits water into oxygen and protons. However, our PSII-I complex does not show any oxygen-evolving activity, suggesting that the OEC is not fully assembled. In mature PSII, the Mn_4CaO_5 cluster is submerged in the complex and additionally capped by the extrinsic subunits PsbO, PsbU and PsbV (Fig. 6a,b). In our PSII-I structure, these subunits are absent, which leaves two parts of the CP43 E loop (residues 320–327 and 397–404) in a flexible conformation, exposing the binding site of the Mn_4CaO_5 cluster to the lumen. There is no strong density feature at this position that would correspond to the fully assembled metal-redox cofactor. Thus, our PSII-I structure provides a model for an immature OEC. Comparing our structure with the high-resolution crystal structure of mature PSII[21] (PDB-ID 3WU2) provides insights into the first steps of OEC biogenesis (Fig. 7).

The D1 C terminus is an essential component for formation of the of OEC, as it provides several charged residues that are responsible for coordination of Cl^- and the Mn_4CaO_5 cluster (Fig. 7a,b,d). The density for these C-terminal residues is weak in our PSII-I map but traceable (Supplementary Fig. 7a), indicating a flexibility that confirms the absence of the OEC. Compared to the mature complex, the last 12 residues of the C-terminal tail of D1 would need to undergo large conformational changes to bring the side chains of Glu333, His337, Asp342 and the Ala344 C terminus into the correct position to coordinate the Mn_4CaO_5 cluster (Supplementary Video 4).

Moreover, we identify a clearly visible density at the position of Cl^-, which is coordinated by Lys317 (D2) and the hydrogen atom of the backbone nitrogen of Glu333 (D1) in mature PSII (Fig. 7b,e). Despite the similar position, the Cl^- is coordinated by the nitrogen atom of the ring of adjacent His332 (D1) in PSII-I (Fig. 7a,e and Supplementary Fig. 7d). Surprisingly, we identified another density in the area where the Mn_4CaO_5 cluster is located in mature PSII (Fig. 7a–c,f and Supplementary Fig. 7c). However, this density is not large enough to reflect the whole cluster. Based on its size and interaction partners (Fig. 7f), it corresponds to one positively charged ion. In the structural context, this ion is probably Mn^{2+} but it could also be Ca^{2+} or any other positively charged ion.

This ion is coordinated by the side chains of D1 Asp170, Glu189 and His332, which are already in similar positions compared to mature PSII. Glu342 and Arg345 (Glu354 and Arg357 in previous publications) of CP43, which are both involved in the second coordination sphere of the Mn_4CaO_5 cluster, are also already prepositioned through the interaction between Arg345 with D1 Asp170 (Fig. 7g). However, there are still substantial conformational changes necessary for the transition from PSII-I to mature PSII, as highlighted in Fig. 7d,g, as well as in Supplementary Video 4. The D1 C-terminal tail must bring the side chains of Glu333, His337 and Asp342, as well as the C terminus of Ala344, into correct alignment to coordinate the Mn_4CaO_5 cluster. In addition, the C-terminal tail of D2 needs to flip towards the D1 C terminus (Fig. 7c, Supplementary Fig. 7b and Supplementary Video 4). In summary, PSII-I is characterized by only one positively charged ion bound instead of the complete Mn_4CaO_5 cluster, resulting in different conformations of the D1 and D2 C-termini compared to the structural model containing a mature Mn_4CaO_5 cluster. However, the PSII-I structure seems to be prepared to accept the Mn_4CaO_5 cluster, as indicated by the similarities in side-chain positioning described above.

Discussion

PSII biogenesis is a complex process that requires the action of specific assembly factors. These auxiliary proteins are not present in the mature complex and interact only transiently with specific subunits or pre-assembled PSII intermediates. Although >20 factors have been identified and assigned to specific transitions, their precise molecular functions in PSII assembly remain elusive in almost all cases. Our study provides detailed molecular insights into the function of PSII assembly factors Psb27, Psb28 and Psb34, which are involved in an important transition before activation of the OEC. The determined binding positions of Psb27 and Psb28, which are two of the most-studied PSII assembly factors, define the binding with atomic resolution, in contrast to previous models[58,59,63–67]. This provides precise information about the molecular function of Psb27 and Psb28 in the assembly process.

Binding of Psb28 and Psb34 to the cytoplasmic side of PSII induces large conformational changes in the D1 D-E loop (Fig. 4), which has been identified previously as an important location for PSII photoinhibition and D1 degradation[68,69]. Structural changes observed in the PSII-I Q_B binding pocket and coordination of the non-haem iron suggest a functional impact on PSII electron transfer to protect the immature complex until water splitting is activated. In particular, D2 Glu241 replacing bicarbonate as a ligand of non-haem iron suggests a regulatory role, as binding of bicarbonate was proposed to tune PSII efficiency by changing the redox potential of (Q_A/Q_A-)[56,57]. As a functional consequence, PSII-I generates less singlet oxygen compared to inactivated PSII (Fig. 5f).

Interestingly, the coordination of the non-haem iron in PSII-I resembles that in non-oxygenic bacterial reaction centres (BRCs)[70] (Supplementary Fig. 6c). In BRCs, the fifth ligand of the non-haem iron is provided by E234 of the M subunit[71] and mutagenesis of this residue induces changes in the free energy gap between the P•+/Q_A•− radical pair[72]. These findings indicate that the environment of the non-haem iron is important for the Q_A redox potential, which influences charge recombination and the formation of reactive oxygen species[56].

The Psb27-bound and -unbound structures do not differ substantially (Fig. 6b), suggesting a rather subtle action in PSII biogenesis. Previous work demonstrated that Psb27 is already bound to free CP43 (ref. [32]), where it might protect free CP43 from degradation or stabilize the E loop in a specific conformation to chaperone the subsequent association with the RC47 complex. This step is crucial for preparing the binding site of the Mn_4CaO_5 cluster, as the CP43 E loop is involved in its second coordination sphere[73]. Moreover, a recent high-speed atomic force microscopy study with spinach PSII has shown fluctuations of the C43 E loop at room temperature after removal of the extrinsic subunits and destruction of the Mn_4CaO_5 cluster[74], which might suggest a role of Psb27 in restraining the scope of conformational fluctuations for further assembly of the OEC. Consistent with this model, recent kinetic analysis of PSII from *Synechocystis* sp. PCC 6803 revealed that Psb27 facilitates photoactivation, probably by stabilization of intermediates with increased accessibly for ions, especially Ca^{2+}, to the site of cluster assembly[62]. However, our data indicate a self-organization of the CP43 E loop, at least in *T. elongatus*. The corresponding structural models for this region are almost identical, independent of binding of Psb27 or the extrinsic proteins, whereas the positions of the D1 and D2 C-termini are clearly different between intermediate and mature PSII. Binding of PsbO seems to be the main driver for these structural differences and its presence interferes with efficient photoactivation, probably by closing the cluster binding site and thus preventing the exchange of ions[62]. The overlap of the Psb27 and PsbO binding sites, although small, might decrease the PsbO affinity and prevent its premature binding, as initially suggested[34,37] or Psb27 may promote binding of PsbO in a different conformation (both proteins can bind simultaneously in a stable complex[49]) that might support photoactivation.

Previous structural studies aimed to obtain mechanistic insights into the dark rearrangement during photoactivation by removing the Mn_4CaO_5 cluster from fully assembled PSII, either by depleting it directly from PSII crystals by chemical treatment[75] or by cryo-EM single-particle analysis in a manganese- and calcium-free buffer[76]. The X-ray structure was indeed missing the Mn_4CaO_5 cluster but the D1 C terminus followed mostly the same trajectory as found in the mature PSII-dimer structure. The authors suggested that the D1 Cterminus might not rearrange during Mn_4CaO_5 cluster assembly. However, the crystal structure was dimeric and still had the extrinsic subunits PsbO, PsbU and PsbV bound. It is known that these subunits are typically not associated with Mn_4CaO_5 cluster-depleted PSII. Thus, the structure might be artificially stabilized by crystal packing forces. The cryo-EM structure, on the other hand, revealed a monomeric PSII that lacks extrinsic subunits and the Mn_4CaO_5 cluster[76]. This structure is like our PSII biogenesis intermediate PSII-I, as PsbY, PsbZ and PsbJ are also missing. The PsbJ subunit is surprising; it is an integral subunit of PSII and should not be easily detached, yet it is missing from this structure, as we deleted it to stabilize our PSII-I complex. These observations might indicate a more specific and regulatory role of PsbJ in PSII biogenesis. Additionally, the D1 C terminus is disordered in this previous cryo-EM structure and the authors suggest that the dark rearrangement involves a transition from a disordered to an ordered state.

Our structure now reveals the fate of the D1 C terminus with the assembly factor Psb27 bound. The D1 C terminus follows a different trajectory compared to the mature PSII. Thus, we provide structural evidence that the slow dark rearrangement involves a conformational change of the D1 C terminus rather than the previously proposed disorder-to-order transition after initial photoactivation[76]. Compared to mature PSII, 12 residues of the D1 C-terminal tail must undergo large conformational changes to bridge the side chains of Glu333, His337 and Asp342, as well as to bring the C terminus of Ala344 to the correct position to coordinate the Mn_4CaO_5 cluster (Figs. 6c and 7d and Supplementary Video 4), which is consistent with previous models[39,77,78]. We also identified a single positively charged ion in our PSII-I structure, coordinated by Asp170, Glu189 and His332 of D1 (Fig. 7f), at the position of the Mn_4CaO_5 cluster of mature PSII. This binding site probably corresponds to the long-sought single HAS, where the first Mn^{2+} binds before the first photoactivation step in OEC biogenesis[41]. The presence of manganese in immature PSII, due to a PratA-facilitated preloading mechanism, has been previously suggested[79]. However, we cannot exclude binding of Ca^{2+}, which was shown to bind with a much lower affinity[77,80], or any other positively charged ion at this position. Nevertheless, Asp170 has been identified as the most critical residue for the HAS[81,82], which supports our hypothesis. Further photoactivation steps occur presumably after cooperative binding of calcium and manganese. The binding of the extrinsic subunit PsbO, potentially after release of Psb27 and maturation of the water-oxidizing complex, is the next step of the PSII assembly line in vivo, which leads to the next unsolved question in PSII biogenesis: what triggers the release of an assembly factor? For Psb27, its detachment might be promoted by the binding of PsbO, as their binding sites partially overlap.

Membrane protein complexes play a fundamental role in bioenergetics to sustain and proliferate life on Earth. They drive the light-to-chemical energy conversion in photosynthetic organisms and are essential for energy supply in heterotrophs. These intricate molecular machines are assembled from numerous single proteins in a spatiotemporally synchronized process that is facilitated by a network of assembly factors. These auxiliary proteins are the key players of nature’s assembly lines. Our PSII-I cryo-EM structure reveals a molecular snapshot of PSII biogenesis and, accompanied by our spectroscopic and biochemical analyses, provides clear mechanistic insights into how three assembly factors (Psb27, Psb28 and Psb34) coordinate the stepwise construction of this powerful catalyst of life.

Methods

Cultivation of *T. elongatus* BP-1.

Cell growth and thylakoid membrane preparation were performed as described previously[83]. In brief, *T. elongatus* mutant strains (Δ*psbJ psbC-TS* and *psb34-TS*) were grown in BG-11 liquid medium inside a 25-l foil fermenter (Bioengineering) at 45 °C, 5% (v/v) CO_2-enriched air bubbling and 50–200 µmol photons m^{-2} s^{-1} white light illumination (depending on the cell density). Cells were harvested at an optical density OD_{680} of ~2 after 5–6 d of cultivation and concentrated to ~0.5 l, using an Amicon DC10 LA hollow fibre system, pelleted (3,500 relative centrifugal force (r.c.f.), 45 min and 25 °C) and resuspended in 150 ml of Buffer D (100 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 500 mM mannitol and 20% (w/v) glycerol). The harvested cells were flash-frozen in liquid nitrogen and stored at −80 °C until further use.

Preparation of *T. elongatus* mutant strains.

T. elongatus Δ*psbJ psbC-TS* was generated on the basis of the previously described strain *T. elongatus* Δ*psbJ*[43] that was transformed with the plasmid pCP43-TS. The plasmid is based on pCP34–10His[34]. The His-tag sequence was exchanged with a Twin-Strep-tag by polymerase chain reaction (PCR) using the primers CP43TS_rev (5′CCCGATATCTTACTTCTCAAATTGCGGATGAGACCACGCAGAACCACCAGAACCACCGCCGCTGCCGCCGCCTTTTTCGAACTGCGGGTGGCTCC3′) and NTCP43 (5′TGCTCTAGAATGAAAACTTTGTCTTCCCAGA3′). The resulting PCR product was ligated back into an empty pCP34–10His backbone using XbaI and EcoRV restriction endonucleases. *T. elongatus* BP-1 cells were transformed as described previously[84]. Mutant colonies were selected by frequent replating onto agar plates with increasing antibiotic concentrations, stopping at 8 μg ml^{-1} of chloramphenicol and 80 μg ml^{-1} of kanamycin. Complete segregation of the mutant was confirmed by PCR with the primers CTCP43DS (5′CCGCTCGAGACCATCCAAGCTTGGCAGCA3′) and NTCP43 (5′TGCTCTAGAATGAAAACTTTGTCTTCCCAGA3′).

T. elongatus psb34-TS was generated by transformation with the plasmid pPsb34-TS. The plasmid DNA was obtained from TwistBioscience. It encodes Psb34 (tsl0063) with a C-terminal Twin-Strep-tag and a kanamycin resistance cassette, flanked by *tsl0063* upstream and downstream regions (900 basepairs each). *T. elongatus* BP-1 cells were transformed[84] and mutant selection took place[34]. Complete segregation of the mutant was verified by PCR. The primers used were *tsl0063*-up-for (5′CATATGGTCTCGCAATTATTTGCCCATGC3′) and *tsl0063*-down-rev (5′GGTACCCCG ACACAGTTGATCACCGC3′).

Purification of PSII assembly intermediates.

Thawed cells were diluted in 100 ml of Buffer A (100 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$ and 10 mM $CaCl_2$) and pelleted again (21,000 r.c.f., 20 min and 4 °C). The pellet was resuspended in 100 ml of Buffer A with 0.2% (w/v) lysozyme and dark incubated for 75–90 min at 37 °C. This was followed by cell disruption by Parr bomb (Parr Instruments) and pelleting (21,000 r.c.f., 20 min and 4 °C). All following steps were performed under green illumination. The pellet was resuspended in 150 ml of Buffer A and pelleted again (21,000 r.c.f., 20 min and 4 °C). This step was repeated three times, with the last resuspension in 80 ml of Buffer B (100 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$ and 500 mM mannitol). The isolated thylakoid membranes were flash-frozen in liquid nitrogen and stored at −80 °C.

Strep-Tactin affinity purification of PsbC-TS and Psb34-TS assembly intermediates were performed under green illumination. Membrane protein extraction was performed as described previously[83], with certain adaptations. Isolated thylakoid membranes were supplemented with 0.05% (w/v) *n*-dodecyl β-maltoside (DDM) (Glycon) and pelleted (21,000 r.c.f., 20 min and 4 °C). The sample was resuspended in extraction buffer (100 mM

Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 1.2% (w/v) DDM, 0.5% (w/v) sodium cholate and 0.01% (w/v) DNase (AppliChem)) to a final chlorophyll concentration of 1 mg ml^{-1} and incubated for 30 min at 20 °C. The solubilized membrane proteins were ultracentrifuged (140,000 r.c.f., 60 min and 4 °C) and NaCl was added to the supernatant to a final concentration of 300 mM.

The supernatant was filtered through a 0.45-μm syringe filter (Sarsted AG) and applied to a 5-ml Strep-Tactin Superflow HC column (IBA Lifesciences), equilibrated in Buffer W (100 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 500 mM mannitol, 300 mM NaCl and 0.03% (w/v) DDM) at a flow rate of 3 ml min^{-1}. The column was washed with Buffer W until a stable baseline (A_{280}) was reached. Strep-tagged protein complexes were eluted by an isocratic elution with Buffer E (100 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 500 mM mannitol, 300 mM NaCl, 2.5 mM desthiobiotin and 0.03% (w/v) DDM). The captured fractions were equilibrated in Buffer F (20 mM MES, pH 6.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 500 mM mannitol and 0.03% (w/v) DDM) with a spin concentrator (Amicon, Ultra-15, nominal molecular weight limit (NMWL) of 100,000 Da), flash-frozen in liquid nitrogen and stored at −80 °C until analysis.

PsbC-TS-containing assembly intermediates were further separated by IEC. Captured elution fractions from the Strep-Tactin affinity purification were loaded onto an anion exchange column (UNO Q6, Biorad) with a flow rate of 4 ml min^{-1}, pre-equilibrated in Buffer F. Protein complexes were eluted by a linear gradient of $MgSO_4$ (0–150 mM) using Buffer G (20 mM MES, pH 6.5, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 500 mM mannitol, 150 mM $MgSO_4$ and 0.03% (w/v) DDM). Fractions containing PSII assembly intermediates were collected, concentrated to 100–10 μM RCs, using a spin concentrator (Amicon, Ultra-15, 100,000 NMWL), aliquoted, flash-frozen in liquid nitrogen and stored at −80 °C until further analysis.

Cryo-EM.

For cryo-EM sample preparation, 4.5 μl of purified protein complexes were applied to glow discharged Quantifoil 2/1 grids, blotted for 3.5 s with force 4 in a Vitrobot Mark III (Thermo Fisher) at 100% humidity and 4 °C, then plunge frozen in liquid ethane, cooled by liquid nitrogen. Cryo-EM data were acquired with a FEI Titan Krios transmission electron microscope using the SerialEM software (v.3–8-0beta8_64; ref. [85]). Video frames were recorded at a nominal magnification of ×22,500 using a K3 direct electron detector (Gatan). The total electron dose of ~55 electrons per $Å^2$ was distributed over 30 frames at a calibrated pixel size of 1.09 Å. Micrographs were recorded in a defocus range of −0.5 to −3.0 μm.

Image processing, classification and refinement.

Cryo-EM micrographs were processed on the fly using the Focus (v.1.1.0) software package[86] if they passed the selection criteria (iciness <1.05, drift $0.4\ Å < x < 70\ Å$, defocus $0.5\ \mu m < x < 5.5\ \mu m$, estimated CTF resolution <6 Å). Micrograph frames were aligned using MotionCor2 (v.1.1.0; ref. [87]) and the contrast transfer function (CTF) for aligned frames was determined using Gctf (v.1.06; ref. [88]). Using Gautomatch (v.056, http://www.mrc-lmb.cam.ac.uk/kzhang/), 693,297 particles were picked template-free on 824 acquired micrographs. Particles were extracted with a pixel box size of 260 using RELION 3.1 (ref. [89]) and imported into Cryosparc 2.3 (ref. [90]). After reference-free two-dimensional classification, 675,123 particles were used for ab initio construction of initial models and subjected to multiple rounds of three-dimensional classification to obtain models with and without Psb28 density. Non-uniform refinement in Cryosparc resulted in models with an estimated resolution of ~3.2 Å. Particles belonging to three-dimensional classes with and without Psb28 (150,090 and 166,411 particles, respectively) were re-extracted in RELION with a pixel box size of 256 and subjected to several rounds of CTF refinement (estimation of anisotropic magnification, fit of per-micrograph defocus and astigmatism and beam tilt estimation) and Bayesian polishing[91]. Both classes were refined using the previously generated starting models. Three-dimensional classification without further alignment using a mask around the Psb27 region separated particles in the Psb28-containing class into distinct classes with and without Psb27 (57,862 and 91,473 particles, respectively). Final refinement of each of the three classes (with Psb27 and Psb28 (PSII-I), with Psb28 but without Psb27 (PSII-I′) and without Psb27 and Psb28 (PSII-M)) resulted in models with global resolutions of 2.94, 2.76 and 2.82 Å, respectively (Gold standard FSC analysis of two independent half-sets at the 0.143 cutoff). Local-resolution and 3DFSC plots (Supplementary Fig. 2) were calculated using RELION and the Remote 3DFSC Processing Server web interface (v.3.0; ref. [92]), respectively.

Atomic model construction.

The 3.6-Å resolution X-ray structure of monomeric PSII from *T. elongatus* with PDB-ID 3KZI[52] was used as initial structural model that was docked as rigid body using UCSF Chimera 1.10.2 (ref. [93]) into the obtained cryo-EM densities for PSII-M and PSII-I. The cofactors that had no corresponding density were removed. The subunit PsbJ was also removed, as it was deleted in the experimental design. By highlighting the still unoccupied parts of the PSII-I density map, we identified densities that led to the structures of Psb27, Psb28 and Psb34.

The 2.4-Å resolution X-ray structures of isolated Psb28 from *T. elongatus* with PDB-ID 3ZPN[94] and the 1.6-Å resolution X-ray structure of isolated Psb27 from *T. elongatus* with PDB-ID 2Y6X[66] were docked as rigid bodies into the unoccupied densities. The 1.6-Å resolution X-ray structure of CyanoQ from *T. elongatus* with PDB-ID 3ZSU[95] does not fit into the density and was therefore not modelled.

As there was no experimentally resolved structural model of Psb34 available, we first used the sequence with UniProt-ID Q8DMP8 to predict structures using the webserver SWISS Model[96] and LOMETS[97]. We also predicted the secondary structure through the meta server Bioinformatics Toolkit[98] and CCTOP[99]. The results of the secondary structure prediction are summarized in Supplementary Table 4. Combining these predictions together with the unassigned cryo-EM density, we used COOT[100] included in CCP4 7.0 to build an initial model of Psb34 that has one α-helix from amino acid numbers 28 to 55.

Model refinement.

The initial model of the complex described above was refined in real space against the cryo-EM density of PSII-I and structural clashes were removed using molecular dynamics flexible fitting (MDFF)[101]. MDFF simulations were prepared in VMD 1.9.4a35 (ref. [102]) using QwikMD[103] and the MDFF plugin. The simulations were carried out with NAMD 2.13 (ref. [104]) using the CHARMM36 force field. Secondary structure, cis peptide and chirality restraints where used during 800 steps of minimization followed by a 40-ps MDFF simulation at 300 K. Due to the restraints used, only conformational changes of side chains and subunit movements compared to the initial structure were identified during the initial MDFF run. We checked the fit to density of the structure by calculating cross-correlation values of the backbone atoms. For PSII-I, we identified residues 217 to 269 from PsbA and residues 467 to 499 from PsbB and PsbZ as main regions where the structural model was not yet in accordance with the density after the initial MDFF run. For these three regions, we used an iterative combination of MDFF with Rosetta 3.8 (refs. [105,106]). Here, we used the optimized strategy as described for model construction of the 26S proteasome[107,108].

To obtain an atomic model that fits the PSII-M density, we used the initial model based on 3KZI described above but without PsbJ, Psb27, Psb28 and Psb34. After the initial MDFF run, the cross-correlation check did not reveal any regions with substantial deviation between model and density. Therefore, no further refinement was necessary. This fast convergence reflects that there are no crucial differences between the PSII-M model and the X-ray structure 3KZI.

To obtain the atomic model that fit the PSII-I′ density, we used the final PSII-I model without Psb27 for MDFF. After the initial MDFF run, the cross-correlation check did not reveal any regions with substantial deviation between model and density. This fast convergence reflects that there are no crucial differences between the PSII-I and PSII-I′ models, except for the presence of the Psb27 subunit.

Lastly, the PSII-M, PSII-I and PSII-I′ models were used to initiate one final round of real-space refinement in Phenix 1.17.1–3660 (ref. [109]) including structure validation checks using MOLPROBITY and EMRinger score.

Visual inspection of structural models and density maps was done with VMD 1.9.4a35 (ref. [102]), UCSF Chimera 1.10.2 (ref. [93]) and PyMol 2.2.0.

Reporting Summary.

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

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank C. König, M. Völkel and R. Oworah-Nkruma for excellent technical assistance, K. Becker for cloning of the pIVEXPsb28His plasmid, B. Erjavec for preparation of the scheme in Fig. 1 and N. Cox for helpful discussion. J.M.S. is grateful to E. Conti for scientific independence and great mentorship and to J. M. Plitzko and W. Baumeister for access to the cryo-EM infrastructure and early career support. M.M.N. is grateful to his mentor M. Rögner for generous support. Financial support was provided by the Max Planck Society, the Helmholtz Zentrum München, the Deutsche Forschungsgemeinschaft (DFG) Research Unit FOR2092 (grant nos. EN 1194/1–1 to B.D.E. and 836/3–2 to M.M.N.), the DFG priority programme 2002 (grant no. 836/4–1 to M.M.N. and grant no. 3542/1–1 to J.D.L.), National Institutes of Health (NIH) grant no. NIH P41-GM104601 (to E.T.) and an Emmy-Noether fellowship (SCHU 3364/1–1 to J.M.S). A.K.-L. was supported by the LabEx Saclay Plant Sciences-SPS (grant no. ANR-10-LABX-0040-SPS) and the French Infrastructure for Integrated Structural Biology (FRISBI; grant no. ANR-10-INSB-05). R.S. gratefully acknowledges support from the DFG (grant nos. INST 213/757–1 FUGG and INST 213/843–1 FUGG).

Data availability

The cryo-EM density maps are deposited in the Electron Microscopy Data Bank under accession numbers EMD-12335 (PSII-M), EMD-12336 (PSII-I) and EMD-12337 (PSII-I′). The atomic models of the cryo-EM structures are deposited in the Worldwide Protein Data Bank under accession numbers 7NHO (PSII-M), 7NHP (PSII-I) and 7NHQ (PSII-I′). The NMR backbone assignments for Psb28 bound to the C-terminal peptide of CP47 are deposited in the Biological Magnetic Resonance Bank under accession code 50747. Protein sequences with the following accession codes were downloaded from the UniProt database (P0A444, Q8DIQ1, Q8DIF8, Q8CM25, Q8DIP0, Q8DIN9, Q8DJ43, Q8DJZ6, Q9F1K9, Q8DIN8, Q8DHA7, Q8DIQ0, Q9F1R6, Q8DJI1, Q8DHJ2, Q8DG60, Q8DLJ8, Q8DMP8). Source data are provided with this paper. Additional data supporting the findings of this manuscript are available from the corresponding authors on reasonable request.

References

1. Hohmann-Marriott MF & Blankenship RE Evolution of photosynthesis. Annu. Rev. Plant Biol 62, 515–548 (2011). [PubMed: 21438681]
2. Sanchez-Baracaldo P & Cardona T On the origin of oxygenic photosynthesis and cyanobacteria. New Phytol 225, 1440–1446 (2020). [PubMed: 31598981]
3. Vinyard DJ, Ananyev GM & Dismukes GC Photosystem II: the reaction center of oxygenic photosynthesis. Annu. Rev. Biochem 82, 577–606 (2013). [PubMed: 23527694]
4. Boekema EJ et al. Supramolecular structure of the photosystem II complex from green plants and cyanobacteria. Proc. Natl Acad. Sci. USA 92, 175–179 (1995). [PubMed: 7816811]
5. Cox N, Pantazis DA & Lubitz W Current understanding of the mechanism of water oxidation in photosystem II and its relation to XFEL data. Annu. Rev. Biochem 89, 795–820 (2020). [PubMed: 32208765]
6. Shen JR The structure of photosystem II and the mechanism of water oxidation in photosynthesis. Annu. Rev. Plant Biol 66, 23–48 (2015). [PubMed: 25746448]
7. Yano J et al. Light-dependent production of dioxygen in photosynthesis. Met. Ions Life Sci 15, 13–43 (2015). [PubMed: 25707465]
8. Ferreira KN, Iverson TM, Maghlaoui K, Barber J & Iwata S Architecture of the photosynthetic oxygen-evolving center. Science 303, 1831–1838 (2004). [PubMed: 14764885]
9. Cardona T, Sedoud A, Cox N & Rutherford AW Charge separation in photosystem II: a comparative and evolutionary overview. Biochim. Biophys. Acta 1817, 26–43 (2012). [PubMed: 21835158]
10. Holzwarth AR et al. Kinetics and mechanism of electron transfer in intact photosystem II and in the isolated reaction center: pheophytin is the primary electron acceptor. Proc. Natl Acad. Sci. USA 103, 6895–6900 (2006). [PubMed: 16641109]

11. Müh F, Glöckner C, Hellmich J & Zouni A Light-induced quinone reduction in photosystem II. Biochim. Biophys. Acta 1817, 44–65 (2012). [PubMed: 21679684]
12. Faller P et al. Rapid formation of the stable tyrosyl radical in photosystem II. Proc. Natl Acad. Sci. USA 98, 14368–14373 (2001). [PubMed: 11762431]
13. Roose JL, Frankel LK, Mummadisetti MP & Bricker TM The extrinsic proteins of photosystem II: update. Planta 243, 889–908 (2016). [PubMed: 26759350]
14. Müh F & Zouni A Structural basis of light-harvesting in the photosystem II core complex. Protein Sci 29, 1090–1119 (2020). [PubMed: 32067287]
15. Shi LX, Hall M, Funk C & Schröder WP Photosystem II, a growing complex: updates on newly discovered components and low molecular mass proteins. Biochim. Biophys. Acta 1817, 13–25 (2012). [PubMed: 21907181]
16. Stewart DH & Brudvig GW Cytochrome b_{559} of photosystem II. Biochim. Biophys. Acta 1367, 63–87 (1998). [PubMed: 9784607]
17. Cox N et al. Electronic structure of the oxygen-evolving complex in photosystem II prior to O-O bond formation. Science 345, 804–808 (2014). [PubMed: 25124437]
18. Kern J et al. Structures of the intermediates of Kok's photosynthetic water oxidation clock. Nature 563, 421–425 (2018). [PubMed: 30405241]
19. Kupitz C et al. Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. Nature 513, 261–265 (2014). [PubMed: 25043005]
20. Suga M et al. An oxyl/oxo mechanism for oxygen–oxygen coupling in PSII revealed by an X-ray free-electron laser. Science 366, 334–338 (2019). [PubMed: 31624207]
21. Umena Y, Kawakami K, Shen JR & Kamiya N Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. Nature 473, 55–60 (2011). [PubMed: 21499260]
22. Krieger-Liszkay A, Fufezan C & Trebst A Singlet oxygen production in photosystem II and related protection mechanism. Photosynth. Res 98, 551–564 (2008). [PubMed: 18780159]
23. Vass I Molecular mechanisms of photodamage in the photosystem II complex. Biochim. Biophys. Acta 1817, 209–217 (2012). [PubMed: 21565163]
24. Shevela D et al. 'Birth defects' of photosystem II make it highly susceptible to photodamage during chloroplast biogenesis. Physiol. Plant 166, 165–180 (2019). [PubMed: 30693529]
25. Heinz S, Liauw P, Nickelsen J & Nowaczyk M Analysis of photosystem II biogenesis in cyanobacteria. Biochim. Biophys. Acta 1857, 274–287 (2016). [PubMed: 26592144]
26. Nickelsen J & Rengstl B Photosystem II assembly: from cyanobacteria to plants. Annu. Rev. Plant Biol. 64, 609–635 (2013). [PubMed: 23451783]
27. Nixon PJ, 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). [PubMed: 20338950]
28. Komenda J et al. Cleavage after residue Ala352 in the C-terminal extension is an early step in the maturation of the D1 subunit of photosystem II in *Synechocystis* PCC 6803. Biochim. Biophys. Acta 1767, 829–837 (2007). [PubMed: 17300742]
29. Komenda J et al. The cyanobacterial homologue of HCF136/YCF48 is a component of an early photosystem II assembly complex and is important for both the efficient assembly and repair of photosystem II in *Synechocystis* sp. PCC 6803. J. Biol. Chem 283, 22390–22399 (2008). [PubMed: 18550538]
30. 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. Philos. Trans. R. Soc. Lond. B 367, 3444–3454 (2012). [PubMed: 23148271]
31. Dobáková 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). [PubMed: 19036835]
32. Komenda J et al. The Psb27 assembly factor binds to the CP43 complex of photosystem II in the cyanobacterium *Synechocystis* sp. PCC 6803. Plant Physiol 158, 476–486 (2012). [PubMed: 22086423]

33. Mamedov F, Nowaczyk MM, Thapper A, Rögner M & Styring S Functional characterization of monomeric photosystem II core preparations from *Thermosynechococcus elongatus* with or without the Psb27 protein. Biochemistry 46, 5542–5551 (2007). [PubMed: 17432833]
34. Nowaczyk MM et al. Psb27, a cyanobacterial lipoprotein, is involved in the repair cycle of photosystem II. Plant Cell 18, 3121–3131 (2006). [PubMed: 17114356]
35. Roose JL & Pakrasi HB The Psb27 protein facilitates manganese cluster assembly in photosystem II. J. Biol. Chem 283, 4044–4050 (2008). [PubMed: 18089572]
36. Bao H & Burnap RL Photoactivation: the light-driven assembly of the water oxidation complex of photosystem II. Front. Plant Sci 7, 578–591 (2016). [PubMed: 27200051]
37. Becker K, Cormann KU & Nowaczyk MM Assembly of the water-oxidizing complex in photosystem II. J. Photochem. Photobiol. B 104, 204–211 (2011). [PubMed: 21382728]
38. Cheniae G & Martin I Photoactivation of the manganese catalyst of O2 evolution. I-Biochemical and kinetic aspects. Biochim. Biophys. Acta 253, 167–181 (1971). [PubMed: 5001676]
39. Dasgupta J, Ananyev GM & Dismukes GC Photoassembly of the water-oxidizing complex in photosystem II. Coord. Chem. Rev 252, 347–360 (2008). [PubMed: 19190725]
40. Radmer R & Cheniae GM Photoactivation of the manganese catalyst of O2 evolution. II. A two-quantum mechanism. Biochim. Biophys. Acta 253, 182–186 (1971). [PubMed: 5001677]
41. Nixon PJ & Diner BA Aspartate 170 of the photosystem II reaction center polypeptide D1 is involved in the assembly of the oxygen-evolving manganese cluster. Biochemistry 31, 942–948 (1992). [PubMed: 1731951]
42. Regel RE et al. Deregulation of electron flow within photosystem II in the absence of the PsbJ protein. J. Biol. Chem 276, 41473–41478 (2001). [PubMed: 11546758]
43. Nowaczyk MM et al. Deletion of *psbJ* leads to accumulation of Psb27-Psb28 photosystem II complexes in *Thermosynechococcus elongatus*. Biochim. Biophys. Acta 1817, 1339–1345 (2012). [PubMed: 22387395]
44. 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). [PubMed: 23084968]
45. Weisz DA et al. Mass spectrometry-based cross-linking study shows that the Psb28 protein binds to cytochrome *b*559 in Photosystem II. Proc. Natl Acad. Sci. USA 114, 2224–2229 (2017). [PubMed: 28193857]
46. Bečková 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). [PubMed: 27530366]
47. Bentley FK, Luo H, Dilbeck P, Burnap RL & Eaton-Rye JJ Effects of inactivating psbM and psbT on photodamage and assembly of photosystem II in *Synechocystis* sp. PCC 6803. Biochemistry 47, 11637–11646 (2008). [PubMed: 18834146]
48. Grasse N et al. Role of novel dimeric photosystem II (PSII)–Psb27 protein complex in PSII repair. J. Biol. Chem 286, 29548–29555 (2011). [PubMed: 21737447]
49. Liu H, Roose JL, Cameron JC & Pakrasi HB 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). [PubMed: 21592967]
50. Weisz DA et al. A novel chlorophyll protein complex in the repair cycle of photosystem II. Proc. Natl Acad. Sci. USA 116, 21907–21913 (2019). [PubMed: 31594847]
51. Mabbitt PD, Wilbanks SM & Eaton-Rye JJ Structure and function of the hydrophilic photosystem II assembly proteins: Psb27, Psb28 and Ycf48. Plant Physiol. Biochem 81, 96–107 (2014). [PubMed: 24656878]
52. Broser M et al. Crystal structure of monomeric photosystem II from *Thermosynechococcus elongatus* at 3.6-Å resolution. J. Biol. Chem 285, 26255–26262 (2010). [PubMed: 20558739]
53. Komenda J & Sobotka R Cyanobacterial high-light-inducible proteins— protectors of chlorophyll-protein synthesis and assembly. Biochim. Biophys. Acta 1857, 288–295 (2016). [PubMed: 26341017]
54. Mulo P et al. Mutagenesis of the D-E loop of photosystem II reaction centre protein D1. Function and assembly of photosystem II. Plant Mol. Biol 33, 1059–1071 (1997). [PubMed: 9154987]

55. Eaton-Rye JJ & Govindjee. Electron transfer through the quinone acceptor complex of photosystem II in bicarbonate-depleted spinach thylakoid membranes as a function of actinic flash number and frequency. Biochim. Biophys. Acta 935, 237–247 (1988).
56. Allen JF & Nield J Redox tuning in photosystem II. Trends Plant Sci 22, 97–99 (2017). [PubMed: 27979715]
57. Brinkert K, De Causmaecker S, Krieger-Liszkay A, Fantuzzi A & Rutherford AW Bicarbonate-induced redox tuning in photosystem II for regulation and protection. Proc. Natl Acad. Sci. USA 113, 12144–12149 (2016). [PubMed: 27791001]
58. Cormann KU, Möller M & Nowaczyk MM Critical assessment of protein cross-linking and molecular docking: an updated model for the interaction between photosystem II and Psb27. Front. Plant Sci 7, 157 (2016). [PubMed: 26925076]
59. Liu H, Huang RY, Chen J, Gross ML & Pakrasi HB 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). [PubMed: 22031695]
60. Wei L et al. LPA19, a Psb27 homolog in *Arabidopsis thaliana*, facilitates D1 protein precursor processing during PSII biogenesis. J. Biol. Chem 285, 21391–21398 (2010). [PubMed: 20444695]
61. Wei X et al. Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. Nature 534, 69–74 (2016). [PubMed: 27251276]
62. Avramov AP, Hwang HJ & Burnap RL The role of Ca2+ and protein scaffolding in the formation of nature's water oxidizing complex. Proc. Natl Acad. Sci. USA 117, 28036–28045 (2020). [PubMed: 33106422]
63. Cormann KU et al. Structure of Psb27 in solution: implications for transient binding to photosystem II during biogenesis and repair. Biochemistry 48, 8768–8770 (2009). [PubMed: 19697957]
64. Fagerlund RD & Eaton-Rye JJ The lipoproteins of cyanobacterial photosystem II. J. Photochem. Photobiol. B 104, 191–203 (2011). [PubMed: 21349737]
65. 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). [PubMed: 23546881]
66. Michoux F et al. Crystal structure of the Psb27 assembly factor at 1.6 Å: implications for binding to photosystem II. Photosynth. Res 110, 169–175 (2012). [PubMed: 22193820]
67. Weisz DA, Gross ML & Pakrasi HB The use of advanced mass spectrometry to dissect the life-cycle of photosystem II. Front. Plant Sci 7, 617 (2016). [PubMed: 27242823]
68. Kettunen R, Tyystjarvi E & Aro EM Degradation pattern of photosystem II reaction center protein D1 in intact leaves. The major photoinhibition-induced cleavage site in D1 polypeptide is located amino terminally of the DE loop. Plant Physiol 111, 1183–1190 (1996). [PubMed: 8756500]
69. Mulo P, Laakso S, Maenpaa P & Aro EM Stepwise photoinhibition of photosystem II. Studies with *Synechocystis* species PCC 6803 mutants with a modified D-E loop of the reaction center polypeptide D1. Plant Physiol 117, 483–490 (1998). [PubMed: 9625701]
70. Stowell MH et al. Light-induced structural changes in photosynthetic reaction center: implications for mechanism of electron–proton transfer. Science 276, 812–816 (1997). [PubMed: 9115209]
71. Wang X et al. Is bicarbonate in photosystem II the equivalent of the glutamate ligand to the iron atom in bacterial reaction centers? Biochim. Biophys. Acta 1100, 1–8 (1992). [PubMed: 1314662]
72. Cheap H et al. M234Glu is a component of the proton sponge in the reaction center from photosynthetic bacteria. Biochim. Biophys. Acta 1787, 1505–1515 (2009). [PubMed: 19632193]
73. Burnap RL D1 protein processing and Mn cluster assembly in light of the emerging photosystem II structure. Phys. Chem. Chem. Phys 6, 4803–4809 (2004).
74. 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). [PubMed: 32551630]
75. Zhang M et al. Structural insights into the light-driven auto-assembly process of the water-oxidizing Mn_4CaO_5-cluster in photosystem II. eLife 6, e26933 (2017). [PubMed: 28718766]
76. Gisriel CJ et al. Cryo-electron microscopy structure of monomeric photosystem II from *Synechocystis* sp. PCC 6803 lacking the water-oxidation complex. Joule 4, 2131–2148 (2020).

77. Kolling DR, Cox N, Ananyev GM, Pace RJ & Dismukes GC What are the oxidation states of manganese required to catalyze photosynthetic water oxidation? Biophys. J 103, 313–322 (2012). [PubMed: 22853909]

78. Zaltsman L, Ananyev GM, Bruntrager E & Dismukes GC Quantitative kinetic model for photoassembly of the photosynthetic water oxidase from its inorganic constituents: requirements for manganese and calcium in the kinetically resolved steps. Biochemistry 36, 8914–8922 (1997). [PubMed: 9220979]

79. Stengel A et al. Initial steps of photosystem II de novo assembly and preloading with manganese take place in biogenesis centers in *Synechocystis*. Plant Cell 24, 660–675 (2012). [PubMed: 22319052]

80. Tyryshkin AM et al. Spectroscopic evidence for Ca^{2+} involvement in the assembly of the Mn_4Ca cluster in the photosynthetic water-oxidizing complex. Biochemistry 45, 12876–12889 (2006). [PubMed: 17042506]

81. Campbell KA et al. Dual-mode EPR detects the initial intermediate in photoassembly of the photosystem II Mn cluster: the influence of amino acid residue 170 of the D1 polypeptide on Mn coordination. J. Am. Chem. Soc 122, 3754–3761 (2000).

82. Cohen RO, Nixon PJ & Diner BA Participation of the C-terminal region of the D1-polypeptide in the first steps in the assembly of the Mn_4Ca cluster of photosystem II. J. Biol. Chem 282, 7209–7218 (2007). [PubMed: 17202088]

83. Kuhl H et al. Towards structural determination of the water-splitting enzyme. Purification, crystallization, and preliminary crystallographic studies of photosystem II from a thermophilic cyanobacterium. J. Biol. Chem 275, 20652–20659 (2000). [PubMed: 10748017]

84. Iwai M, Katoh H, Katayama M & Ikeuchi M Improved genetic transformation of the thermophilic cyanobacterium, *Thermosynechococcus elongatus* BP-1. Plant Cell Physiol 45, 171–175 (2004). [PubMed: 14988487]

85. Mastronarde DN Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol 152, 36–51 (2005). [PubMed: 16182563]

86. Biyani N et al. Focus: the interface between data collection and data processing in cryo-EM. J. Struct. Biol 198, 124–133 (2017). [PubMed: 28344036]

87. Zheng SQ et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332 (2017). [PubMed: 28250466]

88. Zhang K Gctf: real-time CTF determination and correction. J. Struct. Biol 193, 1–12 (2016). [PubMed: 26592709]

89. Scheres SH A Bayesian view on cryo-electron microscopy structure determination. J. Mol. Biol 415, 406–418 (2012). [PubMed: 22100448]

90. Punjani A, Rubinstein JL, Fleet DJ & Brubaker MA cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296 (2017). [PubMed: 28165473]

91. Scheres SH Beam-induced motion correction for sub-megadalton cryo-EM particles. eLife 3, e03665 (2014). [PubMed: 25122622]

92. Tan YZ et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat. Methods 14, 793–796 (2017). [PubMed: 28671674]

93. Pettersen EF et al. UCSF chimera—a visualization system for exploratory research and analysis. J. Comput. Chem 25, 1605–1612 (2004). [PubMed: 15264254]

94. 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). [PubMed: 24126792]

95. Michoux F et al. Crystal structure of CyanoQ from the thermophilic cyanobacterium *Thermosynechococcus elongatus* and detection in isolated photosystem II complexes. Photosynth. Res 122, 57–67 (2014). [PubMed: 24838684]

96. Schwede T, Kopp J, Guex N & Peitsch MC SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 31, 3381–3385 (2003). [PubMed: 12824332]

97. Wu S & Zhang Y LOMETS: a local meta-threading-server for protein structure prediction. Nucleic Acids Res 35, 3375–3382 (2007). [PubMed: 17478507]

98. Zimmermann L et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J. Mol. Biol 430, 2237–2243 (2018). [PubMed: 29258817]

99. Dobson L, Remenyi I & Tusnady GE CCTOP: a consensus constrained TOPology prediction web server. Nucleic Acids Res 43, W408–W412 (2015). [PubMed: 25943549]

100. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of coot. Acta Crystallogr D 66, 486–501 (2010). [PubMed: 20383002]

101. Trabuco LG, Villa E, Schreiner E, Harrison CB & Schulten K Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. Methods 49, 174–180 (2009). [PubMed: 19398010]

102. Humphrey W, Dalke A & Schulten K VMD: visual molecular dynamics. J. Mol. Graphics 14, 33–38 (1996).

103. Ribeiro JV et al. QwikMD—integrative molecular dynamics toolkit for novices and experts. Sci. Rep 6, 26536 (2016). [PubMed: 27216779]

104. Phillips JC et al. Scalable molecular dynamics with NAMD. J. Comput. Chem 26, 1781–1802 (2005). [PubMed: 16222654]

105. Leaver-Fay A et al. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. Methods Enzymol 487, 545–574 (2011). [PubMed: 21187238]

106. Lindert S & McCammon JA Improved cryoEM-guided iterative molecular dynamics—rosetta protein structure refinement protocol for high precision protein structure prediction. J. Chem. Theory Comput 11, 1337–1346 (2015). [PubMed: 25883538]

107. Guo Q et al. In situ structure of neuronal *C9orf72* poly-GA aggregates reveals proteasome recruitment. Cell 172, 696–705 (2018). [PubMed: 29398115]

108. Wehmer M et al. Structural insights into the functional cycle of the ATPase module of the 26S proteasome. Proc. Natl Acad. Sci. USA 114, 1305–1310 (2017). [PubMed: 28115689]

109. Liebschner D et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D 75, 861–877 (2019).

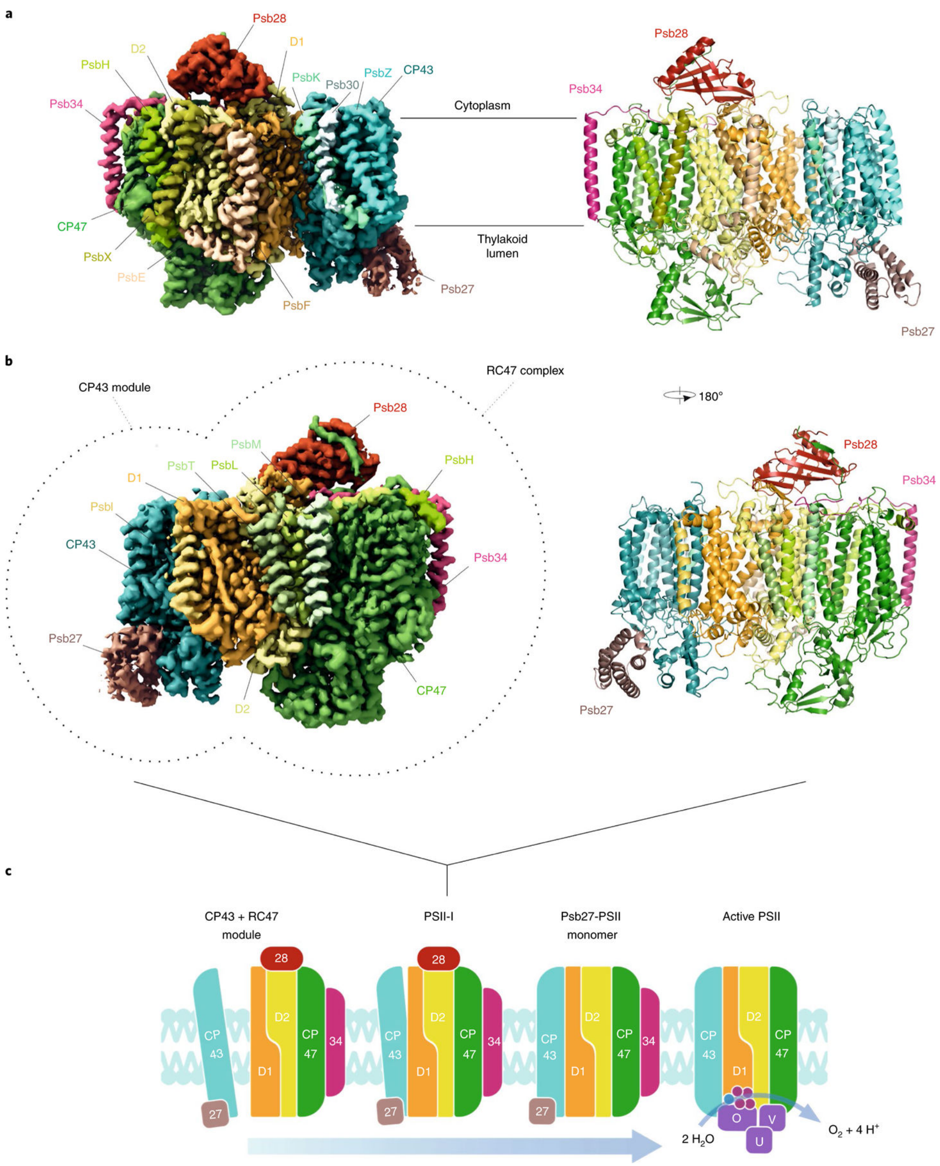

Fig. 1 |. Cryo-EM map of a PSII assembly intermediate (PSII-I) from *T. elongatus*, segmented into subunits.

a, Fifteen PSII subunits and three assembly factors are coloured and labelled (PSII subunits —D1, D2, CP43, CP47, PsbE, PsbF, PsbH, PsbI, PsbK, PsbL, PsbM, PsbT, PsbX, PsbZ and Psb30; assembly factors—Psb27, Psb28 and tsl0063, which we named Psb34) (front view). **b**, Parts of PSII that originate from the CP43 module (comprised of CP43, Psb27, PsbZ, Psb30 and PsbK) and the RC47 complex are indicated by dashed lines (back view). **c**, Schematic model of the PSII assembly process starting with the formation of PSII-I from the CP43 module and RC47. Small PSII subunits were omitted for simplicity.

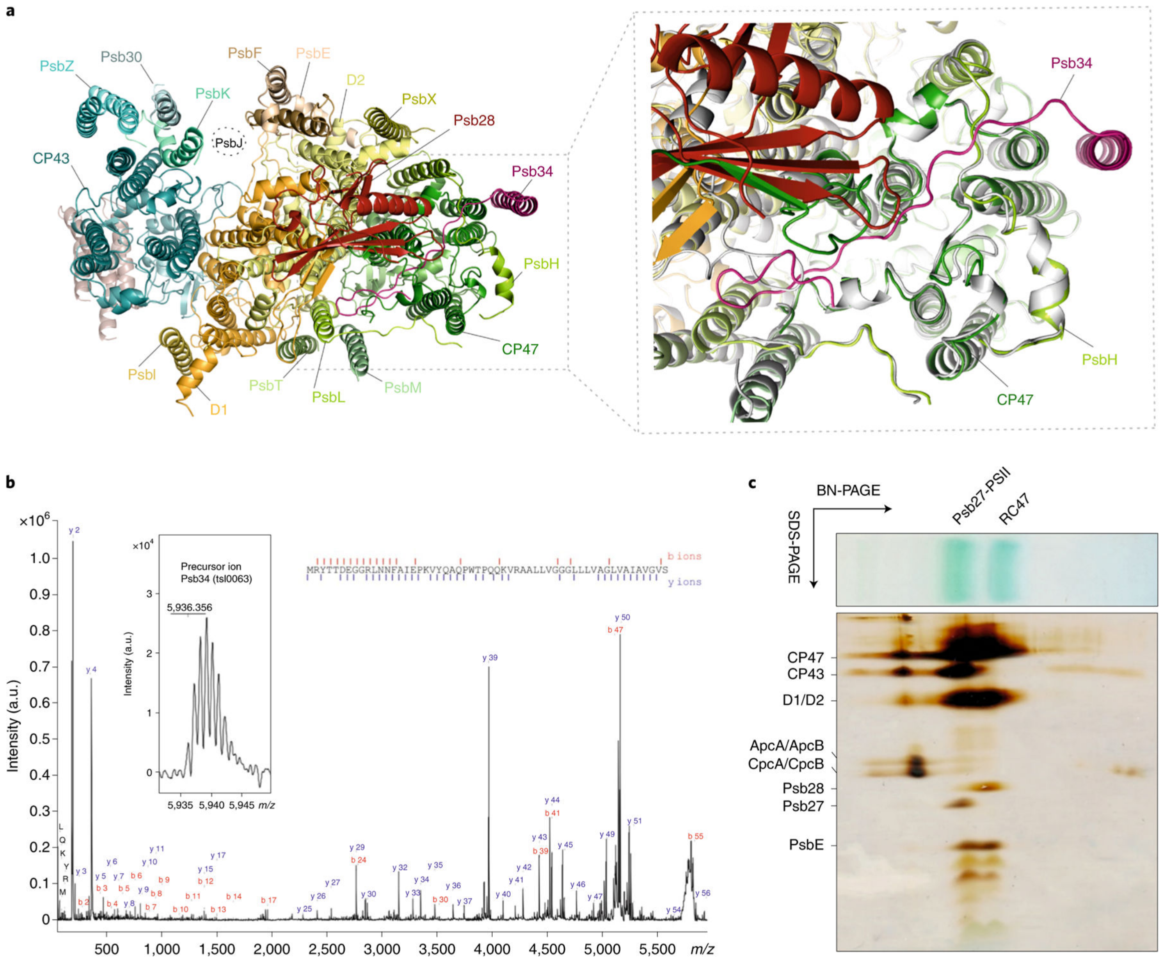

Fig. 2 |. Psb34 binds to RC47 during attachment of the CP43 module.

a, Binding site of Psb34 at CP47, close to PsbH (top view), with extended binding of the Psb34 N terminus along the cytoplasmic PSII surface (dashed box). **b**, Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis of PSII assembly intermediates. Mass spectrum of Psb34 (tsl0063) from the PSII complex (inset) and the fragment spectrum obtained for *m/z* 5,936.356 with annotated b- and y-ion series matching the Psb34 sequence. Observed fragmentation sites are indicated by dashes in the sequence. Mascot score: 171. **c**, Subunit composition of Psb34-PSII assembly intermediates analysed by two-dimensional polyacrylamide gel electrophoresis (PAGE). The results are representative of three independent Psb34-TS preparations. a.u., arbitrary units.

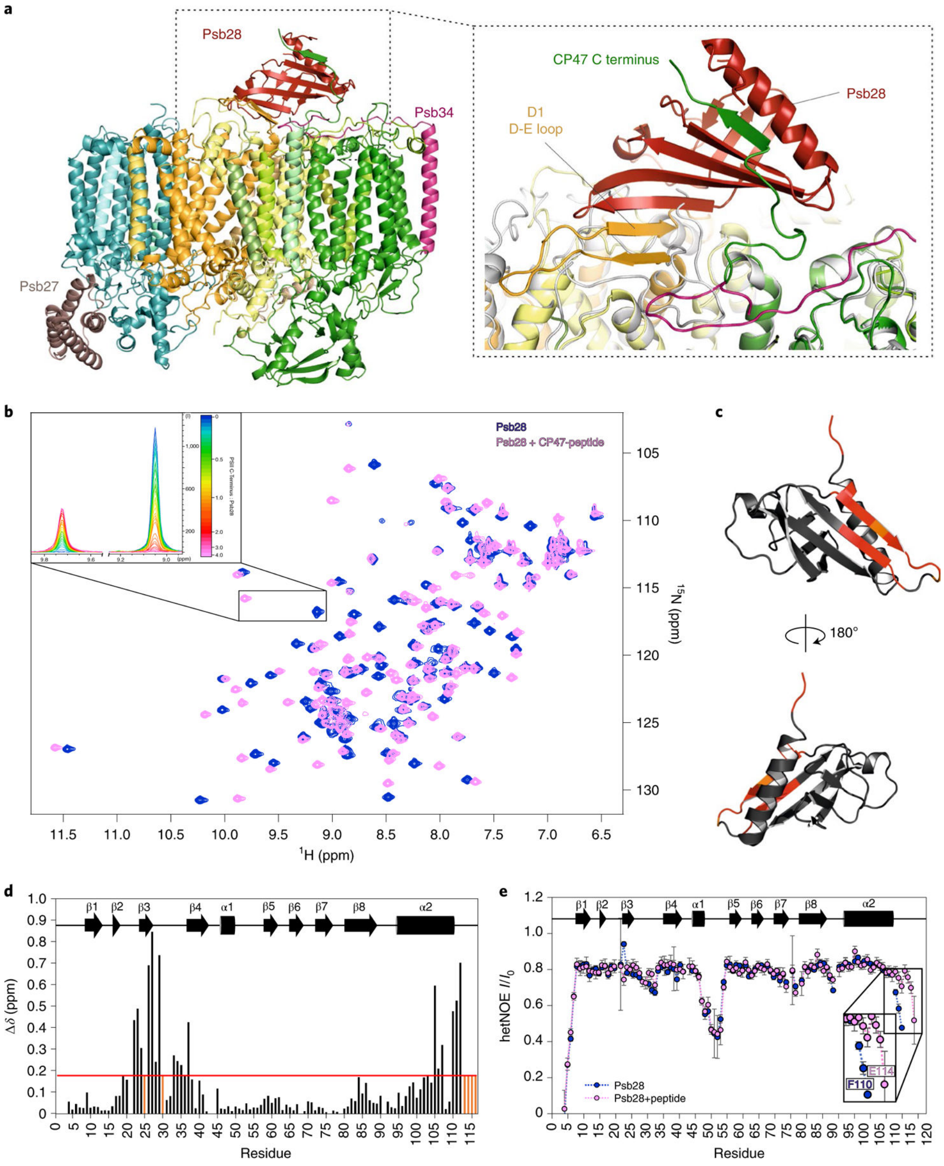

Fig. 3 |. The role of the CP47 C terminus in binding of Psb28.

a, Binding of Psb28 at the cytoplasmic/stromal PSII surface (side view, colours correspond to Fig. 1) and continuation of the central Psb28 β-sheet by the CP47 C terminus and the D-E loop of D1 (dashed box). For comparison, mature monomeric PSII (PDB ID 3KZI) is shown in grey. **b**, Superimposed two-dimensional ^{1}H-^{15}N-HSQC spectra of free Psb28 (blue) and Psb28 bound to the C-terminal peptide of CP47 (magenta). Upper left inset: representation of slow exchange behaviour for the proton amide resonance of T24, ranging from 126.9 to 128.6 ppm in the ^{15}N dimension. **c**, CSPs of >1 s.d. projected onto the model representation of Psb28. **d**, Weighted ^{1}H/^{15}N chemical shift perturbations observed for Psb28 on binding to the CP47 peptide. Red line indicates 1 s.d.; residues that yield resonances only in the complex form are indicated in orange. **e**, Backbone ^{15}N {^{1}H}-heteronuclear NOE of free Psb28 (blue) and Psb28 bound to the C-terminal region of the CP47 peptide (magenta). I/I_0 data are presented as mean values ± s.d. derived from $n = 3$ independent experiments.

Smaller I/I_0 ratios correspond to regions that exhibit dynamics on the pico- to nanosecond timescale.

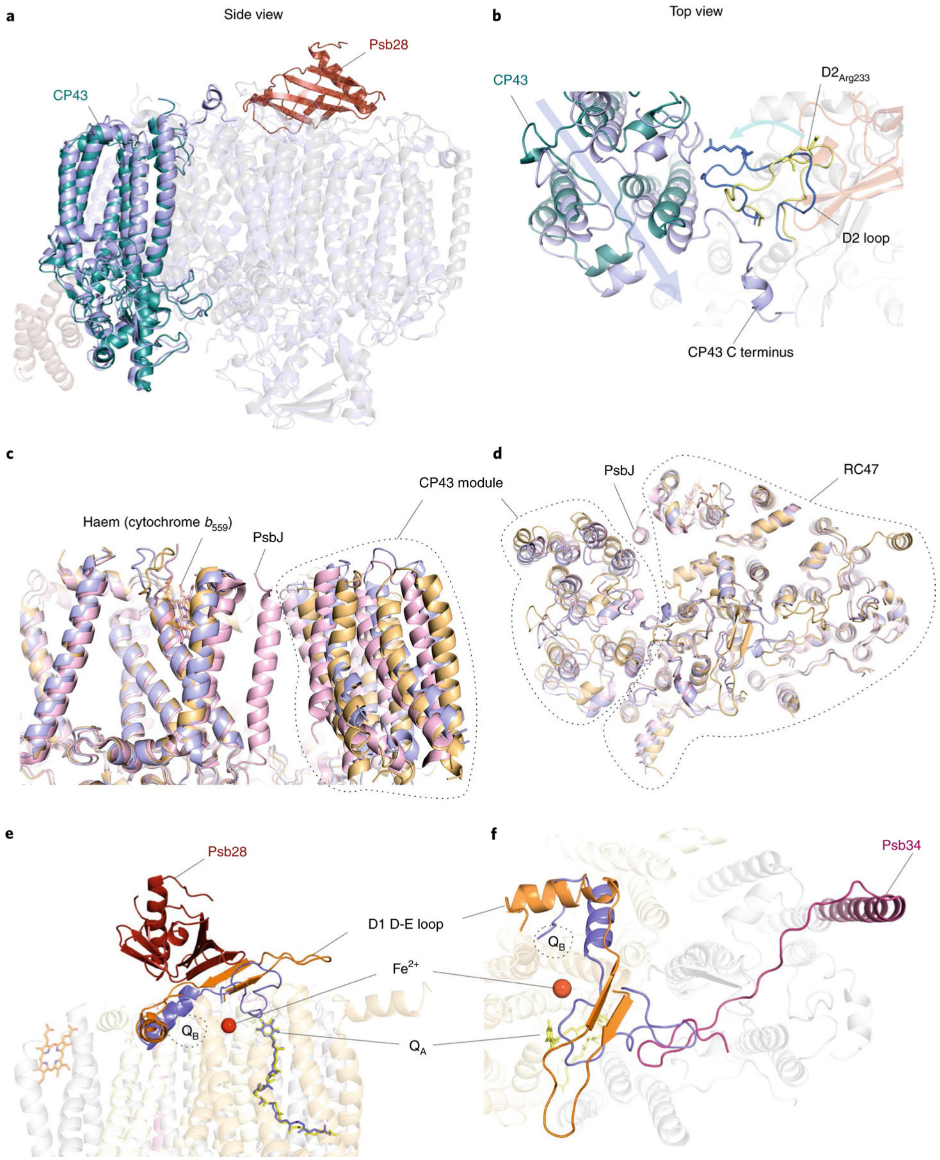

Fig. 4 |. Structural changes of the D1 and D2 D-E loops induced by binding of Psb28 and Psb34.

a, Side view of the CP43 antenna protein in PSII-I (teal) and the PSII-M control (light blue). **b**, Structural changes between PSII-I and the PSII-M control in the cytoplasmic D2 D-E loop (yellow, PSII-I; blue, PSII-M) and attachment of CP43 (teal, PSII-I; light blue, PSII-M control) (top view). Details of the structural changes in the D2 loop are shown in Supplementary Fig. 5a,b. **c**,**d**, Side view (**c**) and top view (**d**) of the PSII-I structure (orange) compared to the PSII-M control (light blue) and mature monomeric PSII (light red, PDB ID 3KZI). **e**,**f**, Side view (**e**) and top view (**f**) of the Psb28-induced structural changes in the D1 D-E loop (orange) and perturbation of the Q_B binding site compared to PSII-M (light blue), which lacks the assembly factors. Q_A is shown in yellow (PSII-I) or light blue (PSII-M), respectively. See Supplementary Fig. 5c–h for enlarged views of the Q_A and Q_B binding sites and the adjacent non-haem iron.

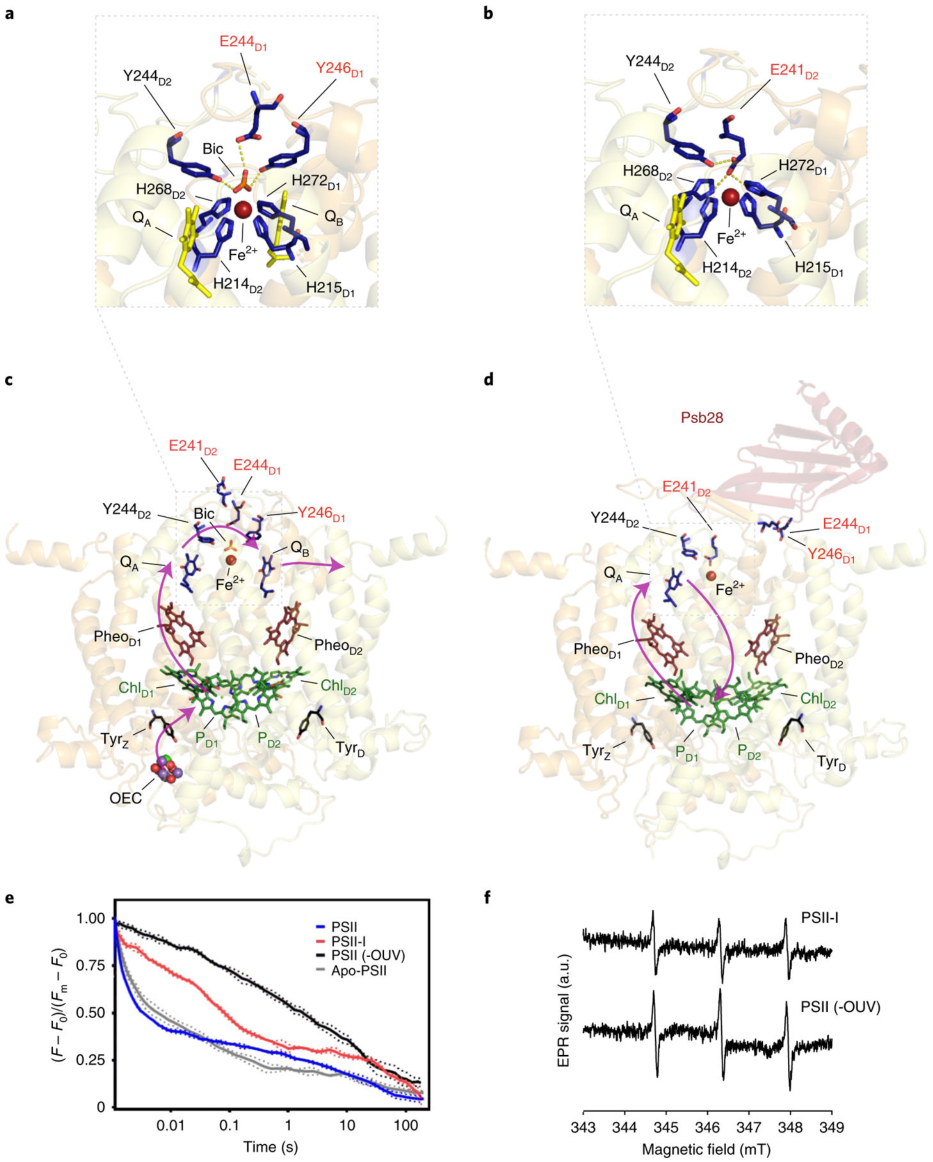

Fig. 5 |. Binding of Psb28 displaces bicarbonate as a ligand of the non-haem iron and protects PSII from damage.

a, The electron transfer from Q_A to Q_B is coordinated by the non-haem iron (Fe^{2+}), with the binding of bicarbonate (Bic) serving as a regulatory mechanism[57] in mature PSII (PDB ID 3WU2). **b**, Binding of Psb28 to the PSII-I assembly intermediate induces a conformational change in the cytoplasmic D2 D-E loop, where the side chain of Glu241 replaces bicarbonate as a ligand of the non-haem iron. The respective fits of the non-haem iron binding sites are shown in (Supplementary Fig. 5e,f). A similar coordination is also found in non-oxygenic bacterial reaction centres[70] (Supplementary Fig. 6c). **c**, Electron transfer (purple arrows) in mature PSII. Light-induced charge separation at the RC chlorophylls (P_{D1}, P_{D2}, Chl_{D1} and Chl_{D2}) leads to electron transfer via pheophytin ($Pheo_{D1}$) and plastoquinone A (Q_A) towards Q_B. The electron gap at the RC is filled by the OEC. **d**, Reoxidation of Q_A^- by direct and safe charge recombination is favoured in the PSII assembly intermediate, as indicated by the purple arrows. **e**, Flash-induced fluorescence decay of PSII. Blue lines represent active PSII and red lines correspond to PSII-I. Black and grey lines represent PSII control samples without a functional OEC (Apo-PSII, hydroxylamine-treated PSII; PSII (-OUV), extrinsic proteins removed). Full lines represent mean values calculated from three independent experiments ($n = 3$), with dotted corridors depicting s.d. **f**, The protective role of Psb28 binding was further confirmed by electron paramagnetic resonance (EPR) spectroscopy using the spin probe TEMPD, which is specific for $^1O_2$, the major reactive oxygen species in PSII generated by triplet chlorophyll (^{3}P).

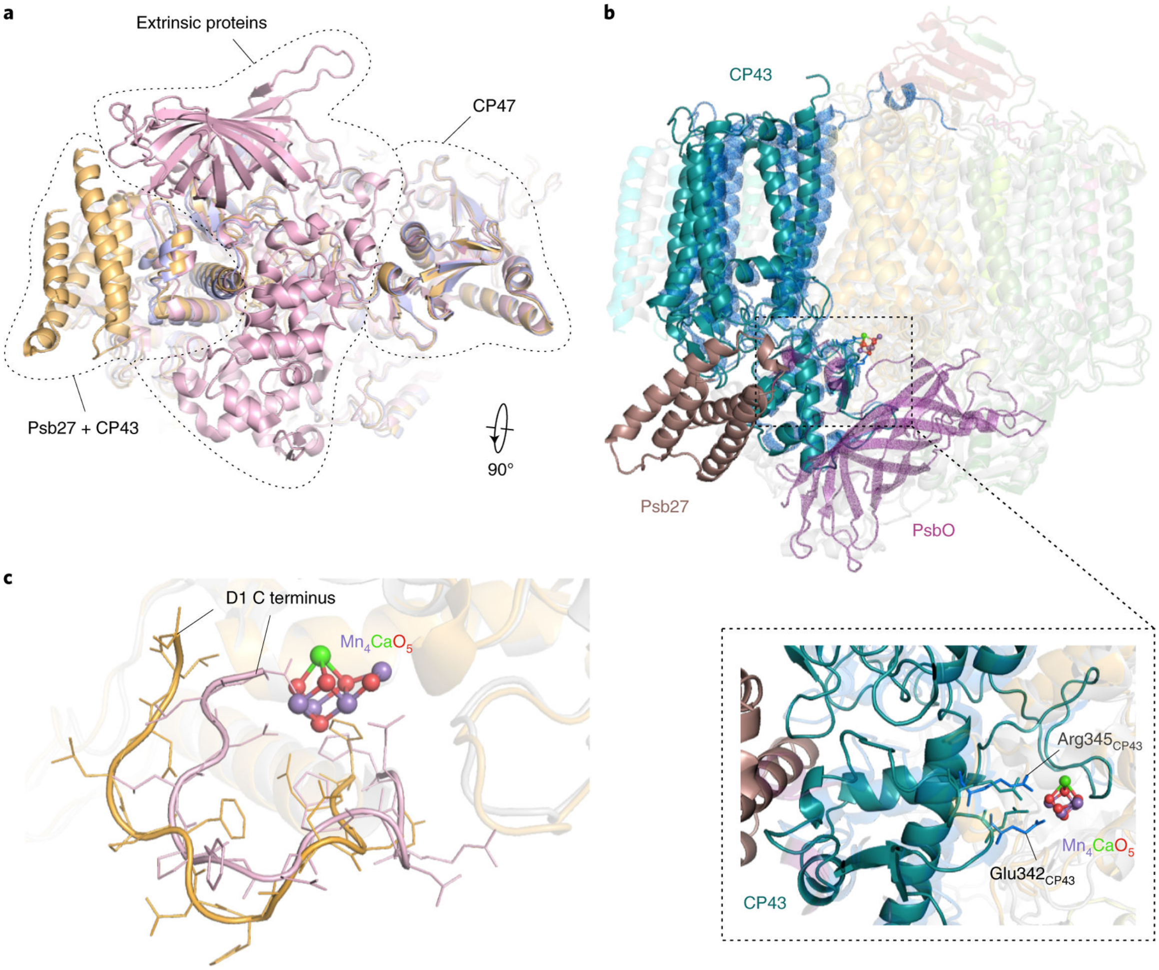

Fig. 6 |. The role of Psb27 in Mn_4CaO_5 cluster assembly.

a, Bottom view of the luminal PSII surface for PSII-I (orange), the PSII-M control (light blue) and mature monomeric PSII (PDB ID 3KZI) (light red). **b**, Side view of CP43 (teal) and Psb27 (brown) in PSII-I, as well as of CP43 (blue) and PsbO (purple) in mature monomeric PSII (PDB ID 3KZI). Dashed box: CP43 E loop with residues Arg345 and Glu342 (shown as sticks), which form the second coordination sphere of the Mn_4CaO_5 cluster. We changed the numbering of CP43 residues due to a corrected N-terminal sequence (www.UniProt.org). The residues correspond to Arg357 and Glu354 in previous publications. The high-resolution structure of the Mn_4CaO_5 cluster is taken from Umena et al.[21] (PDB ID 3WU2). **c**, Position of the D1 C terminus in PSII-I (orange) and mature monomeric PSII (PDB ID 3KZI) (light red).

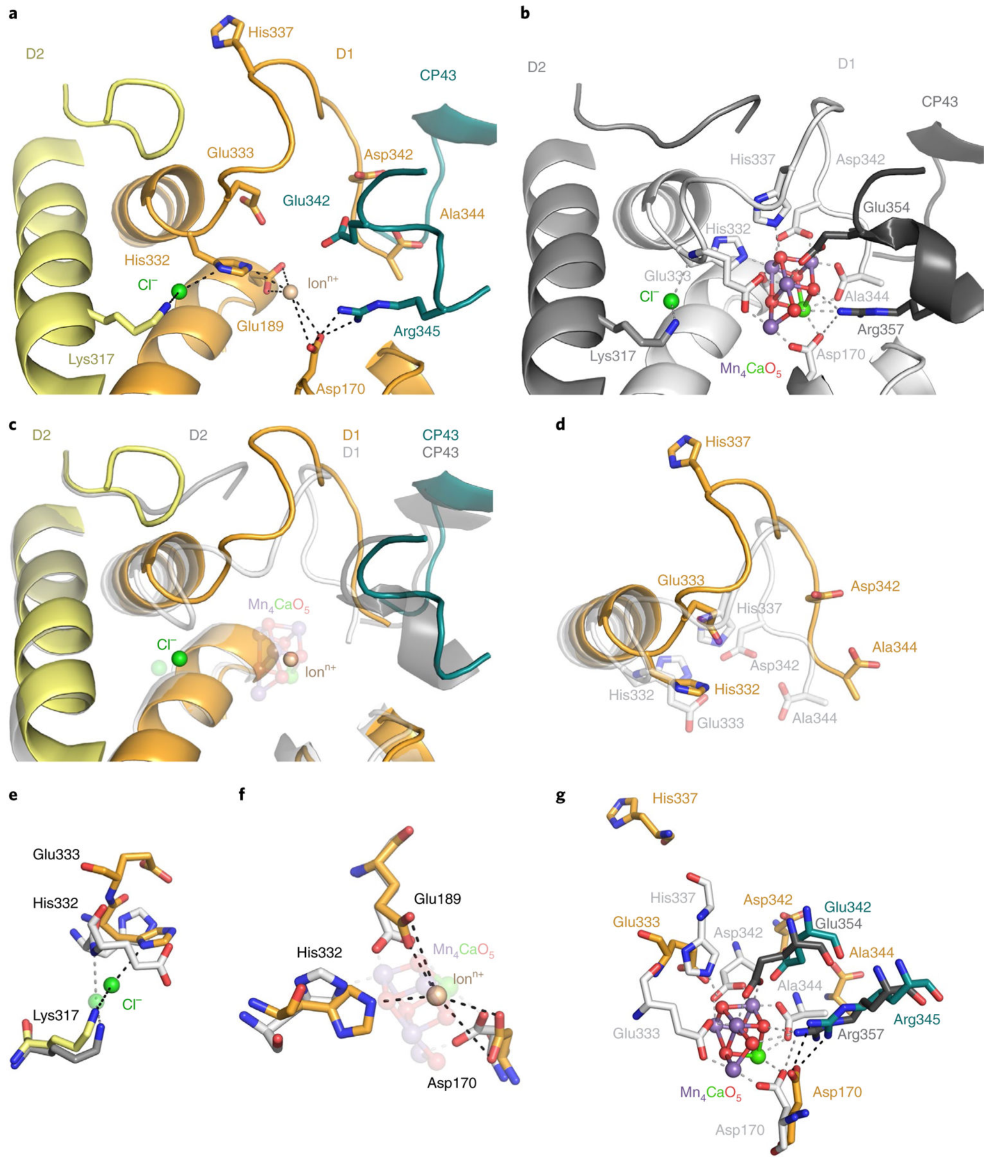

Fig. 7 |. Conformational changes within the active site of the Mn_4CaO_5 cluster. The Mn_4CaO_5 cluster performs the unique water-splitting reaction of PSII.

a, The active site of the Mn_4CaO_5 cluster is resolved within our PSII-I structural model but is not yet oxygen evolving. **b**, Crystal structure of the oxygen evolving, mature PSII (PDB ID 3WU2, resolution 1.9 Å). **c**, Overlay of both structures, illustrating notable differences in the backbone conformation of the D1 and D2 C-terminal tails. **d**, Accompanying side-chain rearrangements of the D1 C-terminus. **e–g**, The Cl⁻ (**e**), ion^{n+} (**f**) and Mn_4CaO_5 (**g**) cluster coordination partners are compared in detail. The validation of the fit to the density of structural details shown here is provided in Supplementary Fig. 7.