How to build a water-splitting machine: structural insights into photosystem II assembly

Jure Zabret, Stefan Bohn, Sandra K. Schuller, Oliver Arnolds, Madeline Möller, Jakob Meier-Credo, Pasqual Liauw, Aaron Chan, Emad Tajkhorshid, Julian D. Langer, Raphael Stoll, Anja Krieger-Liszka, Benjamin D. Engel, Till Rudack, Jan M. Schuller, Marc M. Nowaczyk

Affiliations:
1Department of Plant Biochemistry, Faculty of Biology & Biotechnology, Ruhr University Bochum, 44780 Bochum, Germany.
2Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany.
3Department of Structural Cell Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany.
4SYNMIKRO Research Center, Philipps-University Marburg, 35043 Marburg, Germany.
5Biomolecular Spectroscopy and RUBiospek|NMR, Faculty of Chemistry and Biochemistry, Ruhr-University Bochum, 44780 Bochum, Germany
6Department of Molecular Membrane Biology, Max Planck Institute of Biophysics, 60438 Frankfurt/Main, Germany.
7NIH 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, Illinois
8Max Planck Institute for Brain Research, Max von Laue Strasse 4, 60438 Frankfurt/Main, Germany.
9Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France
10Helmholtz Pioneer Campus, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany.
11Department of Chemistry, Technical University of Munich, Lichtenbergstraße 4, 85748 Garching, Germany.
12Biospectroscopy, Center for Protein Diagnostics (ProDi), Ruhr University Bochum, 44801 Bochum, Germany.
13Department of Biophysics, Faculty of Biology & Biotechnology, Ruhr University Bochum, 44780 Bochum, Germany.

Equal contribution
*corresponding authors: till.rudack@rub.de, jan.schuller@synmikro.uni-marburg.de, marc.m.nowaczyk@rub.de

One Sentence Highlight:
The high-resolution Cryo-EM structure of the photosystem II assembly intermediate PSII-I reveals how nature’s water splitting catalyst is assembled, protected and prepared for photoactivation by help of the three assembly factors Psb27, Psb28 and Psb34.
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 with 2.94 Å resolution. It contains three assembly factors (Psb27, Psb28, 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\textsubscript{B}) and replace bicarbonate with glutamate as a ligand of the non-heme iron, a structural motif found in reaction centers of non-oxygenic photosynthetic bacteria. These results reveal novel 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\textsubscript{4}CaO\textsubscript{5} cluster, which performs the unique water splitting reaction.

Keywords
photosynthesis, photosystem II biogenesis, assembly factors, bicarbonate binding, reactive oxygen species, protection mechanisms, oxygen evolving complex, photoactivation, cryo-EM

Introduction
Photosystem II (PSII) is the only enzyme that catalyzes the light-driven oxidation of water, a thermodynamically demanding reaction that drives photosynthesis, sustaining life on our planet\textsuperscript{1-3}. This multi-subunit 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 (PET). It forms a homodimer with a molecular mass of \textasciitilde500 kDa\textsuperscript{4}, with each monomer composed of at least 20 protein subunits and numerous cofactors, including chlorophylls, quinones, carotenoids, lipids, bicarbonate and the unique Mn\textsubscript{4}CaO\textsubscript{5} cluster\textsuperscript{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\textsuperscript{8}. Light-excitation leads to a charge-separated state in which an electron is transferred from the chlorophyll assembly P\textsubscript{680}\textsuperscript{9} to the nearby pheophytin\textsuperscript{10}. Subsequently, the electron is passed to the bound plastoquinone (Q\textsubscript{A}) and then to the mobile plastoquinone molecule (Q\textsubscript{B}), which leaves the complex after accepting two electrons and two protons\textsuperscript{11}. The electron hole at P\textsubscript{680} is filled by oxidation of an adjacent tyrosine residue (Tyr\textsubscript{Z})\textsuperscript{12} and finally by the oxygen evolving
complex (OEC) that contains the Mn₄CaO₅ cluster. In cyanobacteria, the cluster is shielded at
the luminal side by the three extrinsic proteins, PsbO, PsbU and PsbV, which regulate access
to the OEC by forming a complex network of channels for different substrates and products³.
Light energy is collected and funneled towards P₆₈₀ by the two membrane-intrinsic antenna
proteins CP43 and CP47. These proteins bind most of the chlorophyll molecules and are located
at opposite sides of the D1/D2 heterodimer⁴. Moreover, at least twelve small transmembrane
subunits with one or two transmembrane helices have been identified in PSII⁵, including
cytochrome-b₅₅₉⁶.
Structural and spectroscopic investigations have revealed these aforementioned comprehensive
insights into PSII function⁷-²¹, but we are far from understanding PSII biogenesis with
molecular detail. How nature facilitates the assembly of a multi-subunit, multi-cofactor
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 catalyze the light-driven oxidation of water. This can easily lead to the formation of reactive
oxygen species (e.g., singlet oxygen is produced by triplet chlorophyll in the PSII reaction
center) and subsequent loss of function due to damaged proteins and cofactors²²,²³. 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.
Therefore, PSII biogenesis is not a spontaneous process but rather must be tightly regulated by
the action of assembly factors. Thus far, more than 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 PSII²⁴-²⁶. In
cyanobacteria, PSII biogenesis begins with the formation of the D1/D2 heterodimer reaction
center (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 CtpA²⁷,²⁸. In the next step, the assembly factor Psb28 helps CP47 join
the RC complex to form the RC47 complex, in which iD1 is further processed to its mature
form by CtpA²⁹,³⁰. Almost all ligands of the Mn₄CaO₅ cluster are already present at this stage,
except for those provided by CP43, which comes pre-constructed with assembly factor Psb27
and several small subunits (together called the CP43 module)³¹. 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 PsbV³²-³⁴. Finally, PSII biogenesis completes
with dimerization of two fully assembled monomers and attachment of the soluble
phycobilisome antenna complexes. Interestingly, deletion of pshJ, which encodes a small single
transmembrane helix protein at the entrance of the PSII plastoquinone channel, leads to massive accumulation of an intermediate monomeric PSII complex, which contains both assembly factors Psb27 and Psb28.

Physiological studies of Psb27 and Psb28 deletion strains point towards multifaceted functions. Cyanobacterial mutants lacking Psb28 exhibited slower autotrophic growth, particularly under stress conditions, and limited synthesis of Chl-binding proteins but without decrease in PSII functionality. The Psb28 mutant also exhibited an overall increase in PSII repair and faster recovery from photodamage. Chemical cross-linking combined with mass spectrometry revealed that Psb28 binds to the cytosolic side of CP47 close to cytochrome-b559 and the Qb binding site. Based on this, researchers postulated a protective role for Psb28, whereby it blocks electron transport to the acceptor side of PSII, thereby protecting the RC47 complex from excess photodamage during the assembly process. This hypothesis is strengthened by the observation that Psb28 is also found in PSII repair complexes. 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 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 into 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-EM single particle analysis to describe the first 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 reaction center complex, which precedes incorporation and activation of the Mn4CaO5 cluster. We complement this structural data with spectroscopic analysis, revealing the first detailed insights into the molecular mechanisms of PSII assembly. Our study provides mechanistic answers to three long-standing questions: i) How do assembly factors modulate the structures of PSII subunits to assist biogenesis? ii) How is PSII protected from photodamage during assembly? iii) How is the PSII active site prepared for incorporation of the Mn4CaO5 cluster?
Results

Structure determination of the PSII assembly intermediate (PSI-I)

Stable PSII intermediates were purified from the *T. elongatus* Δ*psbJ* mutant by affinity chromatography using a twin-strep-tag fused to the C-terminus of the CP43 subunit and subsequent ion exchange chromatography (Fig. S1A). 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 (Fig. S1B and 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 (Fig. S2, Table S1). In addition to the protein subunits, we also faithfully assigned all essential non-protein cofactors, including chlorophylls, quinones, carotenoids and lipids, which are also present in the mature PSII complex (Fig. S3). Consistent with previous biochemical studies, the EM density corresponding to the fully assembled, active Mn₄CaO₅ 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 reaction center complex (Fig. 1). This PSII intermediate contains three assembly factors (Psb27, Psb28 and Psb34), as well as almost all the membrane-intrinsic subunits and cofactors found in mature PSII. Psb27 and Psb28 are well-known assembly factors, 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, 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.

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 (RMSD) of the Cα atomic positions between similar subunits of the two complexes is 0.4 Å. Most likely, 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 (PDB-ID 3KZI, 3.6 Å) reveals only minimal differences between both structures, with a Cα RMSD of 1.3 Å, which verifies that the structural changes observed in PSII-I are not caused by the deletion of PsbJ.
Fig. 1: Cryo-EM map of a PSII assembly intermediate (PSII-I) from *T. elongatus*, segmented by subunit. (A) 15 PSII subunits and 3 assembly factors are colored and named (PSII subunits: D1, D2, CP43, CP47, PsbE, PsbF, PsbH, PsbI, PsbK, PsbL, PsbM, PsbT, PsbX, PsbZ and Psb30; assembly factors: Psb27, Psb28 and tsI0063, 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). 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.
Psb34 specifically assists the attachment of the CP43 module to RC47

Our PSII-I structure provides the first 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, indicating a specific function of Psb34 in the attachment of CP43 to RC47 (Fig. 2C). Two distinct PSII intermediates were isolated via pulldown of strep-tagged Psb34: the RC47 complex with bound Psb28 and the subsequent PSII intermediate after attachment of CP43 and Psb27 (Fig. 2C). This observation implies that Psb28 is usually released from the PSII intermediate after attachment of CP43, probably after incorporation of PsbJ, as this trigger is missing in the analyzed ΔpsbJ mutant. Psb34 shows sequence similarity to high-light inducible proteins (HLIPs), which play a role in transient chlorophyll storage and chlorophyll biosynthesis. However, the chlorophyll binding motive is missing in Psb34 (Table S2), suggesting a distinct function for this protein in PSII biogenesis.
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) MALDI-ToF analysis of PSII assembly intermediates. Mass spectrum of Psb34 (tsl0063) from the PSII complex (inset) and the fragment spectrum obtained for m/z 5936.356 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 analyzed by 2D-PAGE.

**Psb28 forms an extended beta hairpin structure that involves the D1 D-E loop and the CP47 C-terminus**

Psb28 binds on the cytosolic faces of the D1 and D2 subunits, directly above the Qb binding site (Fig. 3A), which differs from the position that was previously predicted by mass spectrometry. Its binding induces the formation of an extended beta-hairpin structure that incorporates the central anti-parallel beta-sheet of Psb28, the C-terminus of CP47 and the D1 D-E loop (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 Fig. S4). The CSP measurements indicated significant shifts with a chemical shift difference ($\Delta \delta$) of more than one standard deviation located at strands $\beta$3 and $\beta$4 as well at the C-terminal region of Psb28 (Fig. 3C and D). Upon 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 upon CP47 peptide binding due to creation of an intermolecular $\beta$-sheet (Fig. 3E). 2D-lineshape analysis was performed, yielding a $K_D$ of 53.92 ± 0.41 $\mu$M and a dissociation rate $k_{off}$ of 10.14 ± 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).
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, colors correspond to Fig. 1) and continuation of the central Psb28 beta-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 gray. (B) Superimposed 2D $^1$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 behavior for the proton amide resonance of T24, ranging from 126.9 ppm to 128.6 ppm in the $^{15}$N dimension. (C) CSPs of more than one SD projected onto the model representation of Psb28. (D) Weighted $^1$H/$^{15}$N chemical shift perturbations observed
for Psb28 upon binding to the CP47 peptide. Red line indicates one standard deviation (SD), 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). Smaller $I/I_{0}$ ratios correspond to regions that exhibit dynamics on the pico- to nanosecond timescale.

**Psb28 binding prevents full association of CP43 and distorts the Q_{B} binding pocket**

Binding of Psb28—with support of Psb34—causes major structural perturbations at the PSII acceptor side (Supplementary Movies 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 and 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, Fig. S5A and B). After dissociation of Psb28, the CP43 module undergoes a rigid body rotation where it clicks into place (Fig. 4B-D, Supplementary Movie 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 and 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 and F), as this region coordinates several important PSII cofactors. In functional PSII, after charge separation at the reaction center 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 and F, Fig. S5C and D). The nearby non-heme iron is also already in place in PSII-I (Fig. 4E and F, Fig. S5E and 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 (Fig. S5G). 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 and F, Fig. S5H). Notably, D1 Phe265, which coordinates the head group of Q_{B} in mature PSII, is clearly at a different position^{21} (Supplementary Movie 2).
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 Fig. S5A and B. (C) Side view and (D) top view of the PSII-I structure (orange) compared to the PSII-M control (light blue) and mature monomeric PSII (light red, PDB-ID 3KZI). (E) Side view and (F) top view of the Psb28-induced structural changes in the D1 D-E loop (orange) and perturbation of the Q₈ binding site compared to PSII-M (light blue), which lacks the assembly factors. Q₈ is shown in yellow (PSII-I) or light blue (PSII-M), respectively. See Fig. S5C-H for enlarged views of the Q₈ and Q₉ binding site and the adjacent non-heme iron.
Binding of Psb28 protects PSII during biogenesis

A more detailed analysis of the structural environment close to the Q<sub>A</sub>/Q<sub>B</sub> binding sites revealed differences in the coordination and the hydrogen-bond network of the adjacent non-heme iron, which also indicate functional consequences for PSII electron transfer and charge recombination processes. In mature PSII, the non-heme iron is coordinated by four histidine residues and bicarbonate as the fifth ligand (Fig. 5A and C), whereas in PSII-I, the bicarbonate molecule is replaced by the E241 side-chain of D2 (Fig. 5B and D, Fig. S5E and F, Supplementary Movie 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, Fig. S5E and F, Supplementary Movie 3). Binding of bicarbonate is important for PSII efficiency, as it lowers the redox potential of (Q<sub>A</sub>/Q<sub>A</sub>-) to favor forward electron transport. If charge recombination occurs, the lower redox potential favors indirect charge recombination via P•+/Pheo•-. This back reaction yields triplet chlorophyll and subsequently singlet oxygen, a highly oxidizing species. Changes in the redox potential of (Q<sub>A</sub>/Q<sub>A</sub>-) have been proposed to tune the efficiency of PSII depending on the availability of CO<sub>2</sub> as the final electron acceptor and thereby protect PSII under low CO<sub>2</sub> conditions. Therefore, we used flash-induced variable fluorescence to measure electron transfer in the PSII-I assembly intermediate and inactivated PSII, both of which lack a functional OEC (Fig. 5E, Fig. S6A and B). The fast component is assigned to PSII centers with fast reoxidation of Q<sub>A</sub>- by properly bound Q<sub>B</sub>, the middle component is caused by PSII complexes with inaccurately bound Q<sub>B</sub>, and the slow component is associated with centers that do not contain Q<sub>B</sub> and instead reoxidize Q<sub>A</sub>- through charge recombination with the Mn<sub>4</sub>CaO<sub>5</sub> cluster. Fully functional PSII showed typical Q<sub>A</sub>- reoxidation, which is primarily due to fast electron transfer to Q<sub>B</sub> (Fig. 5E, blue trace). Addition of DCMU blocks electron transfer to Q<sub>B</sub> in active PSII, thereby promoting slow S<sub>2</sub>Q<sub>A</sub>- charge recombination (Fig. S6B, blue trace). Removal of the OEC increases the Q<sub>A</sub> redox potential and promotes very slow Q<sub>A</sub>-Tyr<sup>+</sup>D recombination (Fig. 5E, black trace), which is influenced only minorly by binding of DCMU (Fig. S6B, black trace). PSII-I shows a different behavior (Fig. 5E, Fig. S6B red trace); ~60% of the PSII-I centers decay within 1 s, whereas ~40% decay in PSII (-OEC). To determine whether the replacement of bicarbonate by glutamate affects the energetics of the redox couple Q<sub>A</sub>/Q<sub>A</sub>-, we measured the formation of ¹O<sub>2</sub> by EPR spectroscopy using the spin probe TEMPD. The data clearly show that ¹O<sub>2</sub> formation is reduced by ~30% in PSII-I compared to inactivated PSII (-OEC), which contains bicarbonate (Fig. 5F).
Fig. 5: Binding of Psb28 displaces bicarbonate as a ligand of the non-heme iron and protects PSII from damage. (A) The electron transfer from PQ_A to PQ_B is coordinated by the non-heme iron (Fe^{2+}), with the binding of bicarbonate (Bic) serving as a regulatory mechanism 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-heme iron. The respective fits of the non-heme iron binding sites are shown in Fig. S5E and F. A similar coordination is also found in non-oxygenic bacterial reaction centers (Fig. S6C). (C) Electron transfer (purple arrows) in mature PSII. Light-induced charge separation at the reaction center chlorophylls (P_D1, P_D2, Chl_D1, Chl_D2) leads to electron transfer via pheophytin (Pheo_D1) and plastoquinone A (Q_A) towards Q_B. The electron gap at the reaction center is filled by the oxygen...
Psb27 binds in a remote position 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). In contrast to previously proposed models\(^54,55\), 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 impact on the Mn\(_{4}\)CaO\(_{5}\) cluster binding site (Fig. 6A and B). Instead, Psb27 is bound at a remote position that might be occupied by CyanoQ in the mature complex\(^56\). This localization of Psb27 does not support previous functional models in which bound Psb27 prevents the binding of the extrinsic subunits or plays a direct role in Mn\(_{4}\)CaO\(_{5}\) cluster assembly\(^33,54\). However, Psb27 might stabilize loop E of CP43 in the unassembled state and facilitate its binding to the D1 subunit. This is of particular importance, as loop E of CP43 provides Arg345 and Glu342, two ligands of the Mn\(_{4}\)CaO\(_{5}\) cluster in mature PSII (Fig. 6B, dashed box). Moreover, in the Psb27-bound state (PSII-I), the D1 C-terminus, which is directly involved in coordination of the Mn\(_{4}\)CaO\(_{5}\) cluster\(^21\), is bound away from the cluster (Fig. 6C, Fig. S7). Thus, our PSII-I structure reveals not only how the Psb27 protein binds to CP43 and thus stabilizes it prior to attachment, but also indicates an indirect role for CP43 in maturation of the oxygen evolving cluster that is consistent with functional data from previous studies\(^31,34,40\).
Fig. 6: The role of Psb27 in Mn₄CaO₅ 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 are involved in coordination of the Mn₄CaO₅ 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 Mn₄CaO₅ cluster is taken from Umena et al. 2011 (PDB-ID 3WU2). (C) Position of the D1 C-terminus in PSII-I (orange) and mature monomeric PSII (PDB-ID 3KZI) (light red).

The immature Mn₄CaO₅ cluster binding site of PSII-I contains a single, positively charged ion

The unique Mn₄CaO₅ 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 oxygen evolving complex (OEC) is not fully assembled. In mature PSII, the Mn₄CaO₅ cluster is submerged in the complex and additionally capped by the extrinsic subunits PsbO, PsbU and PsbV (Fig. 6A and 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 Mn₄CaO₅ 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. By comparing our structure with the high-resolution
crystal structure of mature PSII²¹ (PDB-ID 3WU2) provides insights into the first-steps of OEC
biogenesis (Fig. 7).

The D1 C-terminus is one of the key features for the formation of the of OEC, as it provides
several essential charged residues that are responsible for coordination of the chloride ion and
the Mn₄CaO₅ cluster (Fig. 7A, B and D). The density for these C-terminal residues is weak in
our PSII-I map, but traceable (Fig. S7A), 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 significant conformational changes to bring the side chains of Glu333,
His337, Asp342, and the Ala344 C-terminus into the correct position to coordinate the
Mn₄CaO₅ cluster (Supplementary Movie 4).

Moreover, we identify a clearly visible density at the position of the chloride ion, which is
coordinated by Lys317 (D2) and the hydrogen atom of the backbone nitrogen of Glu333 (D1)
in mature PSII (Fig. 7B and 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 and E, Fig. S7D).
Surprisingly, we identified another density in the area where the Mn₄CaO₅ cluster is located in
mature PSII (Fig. 7A-C and F, Fig. S7C). 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 most likely Mn²⁺, but it could also
be Ca²⁺ 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 of CP43, which are both
involved in the Mn₄CaO₅ cluster coordination, are also already pre-positioned through the
interaction between Arg345 with D1 Asp170 (Fig. 7G). However, there are still significant
conformational changes necessary for the transition from PSII-I to mature PSII, as highlighted
in Figure 7D and G, as well as in Supplementary Movie 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 Mn₄CaO₅ cluster. In addition, the C-terminal tail of D2 needs to
flip towards the D1 C-terminus (Fig. 7C, Fig. S7B, Supplementary Movie 4). In summary,
PSII-I is characterized by only one positive charged ion bound instead of the complete
Mn₄CaO₅ cluster, resulting in significantly different conformations of the D1 and D2 C-termini
compared to the structural model containing a mature Mn₄CaO₅ cluster. However, the PSII-I structure seems to be prepared to accept the Mn₄CaO₅ cluster, as indicated by the above described similarities in side chain positioning.

Figure 7: Conformational changes within the active site of the Mn₄CaO₅ cluster. The Mn₄CaO₅ cluster performs PSII’s unique water-splitting reaction. (A) The active site of the Mn₄CaO₅ 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 significant differences in the backbone conformation of the D1 and D2 C-terminal tails. (D) Accompanying side chain rearrangements of the D1 C-terminus. The Cl⁻ (E), Ion⁺ (F) and Mn₄CaO₅ (G) cluster coordination partners are compared in detailed. The validation of the fit to density for the structural details shown here is provided in Figure S7.
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 preassembled PSII intermediates. Although more than 20 factors have been identified and allocated to specific transitions, their precise molecular function in PSII assembly remains elusive in almost all cases. Our study provides the first detailed molecular insights into the function of PSII assembly factors Psb27, Psb28 and Psb34, which are involved in an important transition prior to activation of the OEC. The determined binding positions of Psb27 and Psb28, which are two of the most studied PSII assembly factors, disprove all previous Psb27 and Psb28 binding models and exclude Psb27 from direct involvement in OEC maturation\textsuperscript{54,55,57-61}.

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\textsuperscript{62,63}. Structural changes observed in the PSII-I Q\textsubscript{B} binding pocket and coordination of the non-heme 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 ligand of the non-heme iron by suggests a regulatory role, as binding of bicarbonate was proposed to tune PSII efficiency by changing the redox potential of (Q\textsubscript{A}/Q\textsubscript{A}-)\textsuperscript{48,49}. As a functional consequence, PSII-I generates less singlet oxygen compared to inactive PSII (Fig. 5F).

Interestingly, the coordination of the non-heme iron in PSII-I resembles that in non-oxygenic bacterial reaction centers (BRCs)\textsuperscript{53} (Fig. S6C). In BRCs, the fifth ligand of the non-heme iron is provided by E234 of the M subunit\textsuperscript{64}, and mutagenesis of this residue induces changes in the free energy gap between the P\textsuperscript{+}/Q\textsubscript{A}+ radical pair \textsuperscript{65}. These findings indicate that the environment of the non-heme iron is important for regulation of forward electron transfer to Q\textsubscript{B} versus charge recombination \textsuperscript{48}. Therefore, we speculate that during biogenesis, PSII switches to a mechanism that usually operates in non-oxygenic bacterial reaction centers.

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\textsuperscript{31}, 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\textsubscript{4}CaO\textsubscript{5} cluster, as the CP43 E-loop provides two ligands of the cluster. Further 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\(^{66-69}\). The mechanistic and structural details of this photoactivation process are not yet understood. In the consensus ‘two quantum model’, a single Mn\(^{2+}\) ion bound to the high-affinity site (HAS) 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 unraveling the mechanism of photoactivation.

Previous structural studies aimed to obtain mechanistic insights into the dark-rearrangement by removing the Mn\(_4\)CaO\(_5\) cluster from fully assembled PSII, either by depleting it directly from PSII crystals by chemical treatment\(^{70}\) or by cryo-EM single particle analysis in manganese- and calcium-free buffer\(^{71}\). The X-ray structure was indeed missing the Mn\(_4\)CaO\(_5\) 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 C-terminus might not rearrange during Mn\(_4\)CaO\(_5\) 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\(_4\)CaO\(_5\) 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\(_4\)CaO\(_5\) cluster\(^{71}\). This structure is more similar to 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 and 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 \(^{71}\). Compared to mature PSII, twelve residues of the D1 C-terminal tail must undergo significant conformational changes to bridge the side chains of Glu333, His337 and Asp342, as well as to bring the C-terminus of Ala344 in the correct position to coordinate the Mn\(_4\)CaO\(_5\) cluster (Fig. 6C and 7D, Supplementary Movie 4), which is consistent with previous models \(^{69,72,73}\). 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
$\text{Mn}_4\text{CaO}_5$ cluster of mature PSII. This binding site most likely corresponds to the long-sought single high-affinity site (HAS), where the first Mn$^{2+}$ binds prior to the first photoactivation step in OEC biogenesis\cite{74}. However, we cannot exclude binding of Ca$^{2+}$, which was shown to bind with a much lower affinity\cite{72,75}, or any other positive charged ion at this position. Nevertheless, Asp170 has been identified as the most critical residue for the HAS\cite{76,77}, 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 WOC, is the next step of the PSII assembly line \textit{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 highly complex 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 the first 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 *Thermosynechococcus elongatus* BP-1

Cell growth and thylakoid membrane preparation were performed as described previously\(^7\). In brief, *T. elongatus* mutant strains (ΔpsbJ psbC-TS and psb34-TS) were grown in BG-11 liquid medium inside a 25-liter 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 OD\(_{680}\) of ~ 2 after 5-6 days of cultivation and concentrated to ~ 0.5 l using an Amicon DC10 LA hollow fiber system, pelleted (3500 rcf, 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

*Thermosynechococcus elongatus* ΔpsbJ psbC-TS was generated based on the previously described strain *T. elongatus* ΔpsbJ\(^3\) that was transformed with the plasmid pCP43-TS. The plasmid is based on pCP34-10His\(^3\). The His-tag sequence was exchanged with TwinStrep-tag by PCR using the primers CP43TS_rev (5´CCCGATATCTTACTTCTCAAATTGGAGACCACGAGAACCACAAGGCTGCCGCCGCCTTTTTCGAACTGGCGGGTGGCTCC 3´) and NTCP43 (5´ TGCTCTAGAATGAAAACTTTGTCTTCCCAGACTGGCGGGTGGCTCC 3´). 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\(^7\). Mutant colonies were selected by frequent re-plating onto agar plates with increasing antibiotic concentrations, stopping at 8 µg/ml of chloramphenicol and 80 µg/ml of kanamycin. Complete segregation of the mutant was confirmed by PCR with the primers CTCP43DS (5´ CCGCTCGAGATTTTCTTGGGCCAGCA 3´) and NTCP43 (5´ TGCTCTAGAATGAAAAACTTTGTCTTCCCAGA 3´). *T. elongatus* psb34-TS was generated by transformation with the plasmid pPsb34-TS. The plasmid DNA was obtained from TwistBioscience. It consisted of psb34 (tsl0063) with a C-terminal TwinStrep-tag and a kanamycin resistance cassette, flanked by tsl0063-upstream and downstream regions (900 bp each). *T. elongatus* BP-1 cells were transformed\(^7\) and mutant selection took place\(^3\). Complete segregation of the mutant was verified by PCR. The primers used were tsl0063-up-for (5´ CATATGGTTGATCCATCCATCGTCTGTC 3´) and tsl0063-down-rev (5´ GGTAACCCGACACAGTTGATCCATCGTCTGTC 3´).
Purification of photosystem II 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 rcf, 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 Company) and pelleting (21 000 rcf, 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 rcf, 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 thylakoids 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$^{78}$, with certain adaptations. Thylakoid membranes were supplemented with 0.05% (w/v) n-Dodecyl β-maltoside (DDM) (Glycon) and pelleted (21000 rcf, 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) to a final chlorophyll concentration of 1 mg/ml and incubated for 30 min at 20 °C. The solubilized membrane proteins were ultra-centrifugated (140000 rcf, 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 filter 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 flowrate of 3 ml/min. 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, 100000 NMWL), flash-frozen in liquid nitrogen and stored at -80 °C until analysis.

PsbC-TS containing assembly intermediates were further separated by ion exchange chromatography (IEC). Captured elution fraction from the Strep-Tactin-affinity purification were loaded onto an anion exchange column (UNO Q-6, Biorad) with a flowrate of 4 ml/min, pre-equilibrated in Buffer F. Protein complexes were eluted by a liner gradient of MgSO$_4$ (0-
150 mM) using Buffer G (20 mM MES, pH 6.5, 10 mM MgCl₂, 10 mM CaCl₂, 500 mM mannitol, 150 mM MgSO₄ and 0.03% (w/v) DDM). Fractions containing PSII assembly intermediates were collected, concentrated to 100 – 10 µM reaction centers, using a spin concentrator (Amicon, Ultra – 15, 100 000 NMWL), aliquoted, flash frozen in liquid nitrogen and stored at -80 °C until further analysis.

Protein Expression and Purification of Psb28

The Psb28 expression plasmid was constructed by first amplifying psb28 from *T. elongatus* genomic DNA, using primers TeloPsb28for (5´ GGAATTCCATATGGTGCAATGGCAGAAATTC 3´) and TeloPsb28rev (5´ CGAATTCCCCGGGAGAGTTCTCAGACTTCTG 3´).

Next, the amplified DNA was cloned into pIVEX2.3d using NdeI/SmaI to obtain pIVEXPsb28His. Expression and purification of ¹⁵N-labelled Psb28 was carried out as described previously with certain adaptations. Overnight starter cultures were grown on agar plates at 37 °C, supplemented with 1 % (w/v) glucose and 100 µg/ml ampicillin. The cell material was then resuspended in 2 ml of M9 media, and this was used to inoculate 500 ml of M9 media with ¹⁵NH₄Cl as the only nitrogen source. Cultures were incubated at 37 °C under vigorous shaking and at an OD₆₀₀ of 0.6, and expression was induced with addition of isopropylthiogalactoside to a final concentration of 0.5 mM. After overnight incubation the ¹⁵N-labbled Psb28 was isolated and purified as described for his-tagged CyanoP. The purity and integrity of the protein samples were checked by SDS-PAGE (data not shown).

Polyacrylamide Gel Electrophoresis

Blue-native PAGE was used to assess the oligomeric state of the isolated PSII assembly intermediates. Separation of protein complexes was carried out across a linear gradient of polyacrylamide (acrylamide-bisacrylamide, 32:1) from 3.2 to 16% (w/v) in the separating gel. This was overlaid with a 3% (w/v) polyacrylamide (acrylamide-bisacrylamide, 32:1) sample gel. The gels were loaded with 40 µg of protein per lane. Electrophoresis was performed at 4 °C in a Mini-PROTEAN Tetra System (BioRad) at 100 V for 30 min with Blue Cathode Buffer (15 mM BisTris-HCl, pH 7.0, 50 mM Tricine and 0.002% (w/v) Coomassie Brilliant Blue 250) and at 170 V for an additional 90 min with Cathode Buffer (15 mM BisTris-HCl, pH 7.0 and 50 mM Tricine). The anode buffer was composed of 50 mM BisTris-HCl at pH 7.

Subunit composition was investigated by SDS-PAGE. Separation of polypeptide chains took place on a 19% (w/v) polyacrylamide gel (acrylamide-bisacrylamide, 37.5:1), containing 9 M urea and 4% (w/v) glycerol. The gel was loaded with 40 µg of denatured protein complex per lane. The gels ran at 4 °C in a Mini-PROTEAN Tetra System (BioRad) at 35 mA per gel for 60
min. Fixation and visualization of polypeptide chains was performed with Coomassie Staining Solution (45% (w/v) isopropanol, 10% (w/v) acetate and 0.2% (w/v) Coomassie Brilliant Blue 250).

**Mass spectrometry analysis**

PSII-I complexes were purified and desalted using Isolute C18 SPE cartridges (Biotage, Sweden). The columns were first washed and equilibrated, the sample diluted in 0.1% trifluoroacetic acid (TFA) and loaded onto the column. After washing with 2 ml 0.1% TFA, the proteins were eluted with 500 µl 80% acetonitrile (ACN), 20% water. The organic fraction was lyophilized in a vacuum concentrator (Eppendorf, Germany), reconstituted in 0.1% TFA and mixed in a 1:1 ratio with HCCA matrix solution (HCCA (alpha-cyano-4-hydroxycinnamic acid) saturated in 50% ACN, 50% water and supplemented with 0.1% TFA). Subsequently, 1 µl aliquots of the mixture were deposited on a ground steel MALDI target and allowed to dry and crystallize at ambient conditions.

MS and MS/MS spectra were acquired on a prototype rapifleX MALDI-TOF/TOF (Bruker Daltonics, Germany) in positive ion mode. The Compass 2.0 (Bruker Daltonics, Germany) software suite was used for spectra acquisition and processing (baseline subtraction, smoothing, peak picking), a local Mascot server (version 2.3, Matrixscience, UK) was used for database searches against the *T. elongatus* proteome (UniProt, retrieved 4/2019) and BioTools 3.2 (Bruker Daltonics) was used for manual spectrum interpretation, de novo sequencing and peak annotation.

**Flash-induced fluorescence decay measurements**

Flash-induced fluorescence decay was measured on a FL3500 Dual-Modulation Kinetic Fluorometer (PSI Photon Systems Instruments). Reaction centers were exited with 625 nm LEDs for both actinic (50 µs) and measuring flashes. The first data point was collected 80 µs after the actinic flash. Data points were collected from 80 µs to 50 or 200 s after the actinic flash for measurements with whole cells and isolated PSII, respectively. 10 data points were collected per logarithmic decade. Assays were performed at room temperature in the presence and absence of 20, 100, 200 or 400 µM 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) with 5 min of dark incubation prior to measurement. Assays with isolated PSII complexes were carried out with 200 nM reaction centers in activity buffer (100 mM KCl, 20 mM MES-KOH, pH 6.5, 10 mM MgCl₂, 10 mM CaCl₂ and 0.03% (w/v) DDM).
Typically, NMR samples contained up to 1 mM of protein in 20 mM Tris/HCl pH 8, 10% D2O, 0.02% NaN3, and DSS. NMR spectra were acquired at 298 K on Bruker DRX 600 and AVANCE III HD 700 spectrometers. Backbone assignments for the free form of Psb28 were obtained from three-dimensional HNCA, HNCO, HN(CO)CACB, HNCACB, and HN(CA)CO spectra. Side-chain assignments were obtained from three-dimensional 1H-15N-HNHA, 1H-13C-HCCH-TOCSY, 1H-15N-HSQC-TOCSY, 1H-15N-HSQC-NOESY, 1H-13C-HSQC-NOESY, and aromatic 1H-13C-HSQC-NOESY spectra. Spectra were processed with NMRPipe and analysed with CcpNmr Analysis. NMR experiments for the complex form of Psb28 and the C-terminal peptide of CP47 were carried out on a Bruker AVANCE III HD 700 spectrometer at 298 K in 20 mM Tris/HCl pH8, 10% D2O, 0.02% NaN3, and DSS. For the backbone assignments of the complex, a 1 mM sample of [U-15N,13C]-enriched Psb28 was mixed with a three-fold excess of peptide. Three-dimensional (3D) HNCA, HNCO, HN(CO)CACB, HNCACB, and HN(CA)CO were recorded with 16 scans and 25% non-uniform sampling (NUS). (H)CC(CO)NH and H(CCCO)NH spectra were recorded with 64 scans and 25% NUS. HNHA and 1H-15N-HSQC-NOESY spectra were recorded with 32 scans and 25% NUS as well as traditional acquisition schemes, respectively. The mixing time for NOESY spectra was set to 120 ms. In addition, heteronuclear two-dimensional 15N {1H}-NOE data were recorded in order to extract pico- to nanosecond dynamics. The titrations were carried out by adding increasing amounts of a peptide stock solution to the NMR sample containing 0.138 mM of protein and two-dimensional 1H-15N-HSQC spectra were recorded after thorough mixing of the Psb28-CP47 carboxyterminal peptide solution. Spectra were processed with NMRPipe and Psb28 ligand affinity calculations based on two-dimensional lineshape analysis were carried out using the TITAN software package.

Synthetic Peptide

The carboxy-terminal peptide from residues 480-499 of CP47, which comprises the sequence SGIDPELSPELSPEQVEWGFYQKV and includes an acetylated amino-terminus, was purchased from JPT Peptide Technologies GmbH, Germany. Peptide stock solutions of at least 6.07 mM for titration experiments were prepared by dissolving the peptide in 20 mM Tris/HCl pH8.

Removal of the PSII oxygen evolving cluster

PSII without functional oxygen evolving cluster (OEC) was prepared by isolating PSII as described by Grasse et al., followed by removal of the extrinsic subunits according to Shen and Inoue, with modifications. PSII was applied to a size exclusion column (Superdex 75
10/300 GL, GE Healthcare) pre-equilibrated in CaCl₂ buffer (10 mM MgCl₂, 20 mM MES-NaOH, pH 6.5, 1 M CaCl₂, 0.03% (w/v) DDM). PSII particles lacking the extrinsic subunits and the Mn₃CaO₅ cluster, which were eluted in the void volume, were collected and the buffer was exchanged to activity buffer (100 mM KCl, 20 mM MES-KOH, pH 6.5, 10 mM MgCl₂, 10 mM CaCl₂ and 0.03% (w/v) DDM) using a spin concentrator (Amicon, Ultra – 15, 100 000 NMWL).

Detection of singlet oxygen by the room temperature EPR spectroscopy

Singlet oxygen was trapped using the water-soluble spin-probe 2,2,6,6-tetramethyl-4-piperidone (TEMPD) hydrochloride¹⁰⁵ and measured with ESR300 (Bruker Biospin, Rheinstetten, Germany). Samples (30 µg chl ml⁻¹) were illuminated for 1 min with red light (RG 630) at 450 µmol quanta m⁻² s⁻¹ in 0.5 M mannitol, 10 mM CaCl₂, 10 mM MgCl₂, 20 mM MES at pH 6.5. Spectra were recorded using a flat cell containing 200 µl sample. The microwave power was 9.77 GHz and 14.07 mW with a modulation frequency of 86 kHz and amplitude of 1.0 G. Each spectrum is an average of 8 scans, each with a sweep time of 10.5 s.

Cryo-electron microscopy

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 was acquired with a FEI Titan Krios transmission electron microscope using the SerialEM software¹⁰⁶. Movie frames were recorded at a nominal magnification of 22,500x using a K3 direct electron detector (Gatan). The total electron dose of ~55 electrons per Å² was distributed over 30 frames at a calibrated physical 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 software package¹⁰⁷ if they passed the selection criteria (iciness < 1.05, drift 0.4 Å < x < 70 Å, defocus 0.5 um < x < 5.5 um, estimated CTF resolution < 6 Å). Micrograph frames were aligned using MotionCor2¹⁰⁸ and the contrast transfer function (CTF) for aligned frames was determined using Gctf¹⁰⁹. Using Gautomatch (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 3.1¹¹⁰ and imported into Cryosparc 2.3¹¹¹. After reference-free 2D classification, 675,123 particles were used for ab initio construction of initial models and subjected to multiple rounds of 3D 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 3D classes with and without Psb28 (150,090 and 166,411 particles, respectively) were reextracted 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. Both classes were refined using the previously generated starting models. 3D 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 Å, 2.76 Å and 2.82 Å, respectively (Gold standard FSC analysis of two independent half-sets at the 0.143 cutoff). Local-resolution and 3D-FSC plots (Extended Data Fig. 2) were calculated using RELION and the “Remote 3DFSC Processing Server” web interface, respectively.

Atomic model construction

The 3.6 Å resolution X-ray structure of monomeric PSII from *T. elongatus* with PDB-ID 3KZI was used as initial structural model that was docked as rigid body using Chimera 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 lead 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 and the 1.6 Å resolution X-ray structure of isolated Psb27 from *T. elongatus* with PDB-ID 2Y6X 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 does not fit into the density and was therefore not modeled.

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 and LOMETS. We also predicted the secondary structure through the meta server Bioinformatics Toolkit and CCTOP. The results of the secondary structure prediction are summarized in Table S4. Combining these predictions together with the unassigned cryo-EM density, we used COOT to build an initial model of Psb34 that has one α-helix from amino acid number 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)\textsuperscript{122}. MDFF simulations were prepared in VMD 1.9.4a35\textsuperscript{123} using QwikMD\textsuperscript{124} and the MDFF plugin. The simulations were carried out with NAMD 2.13\textsuperscript{125} employing the CHARMm36 force field. Secondary structure, cis peptide and chirality restraints where employed during 800 steps of minimization followed by a 40 ps MDFF simulation at 300K. Due to the employed restraints, only conformational changes of side chains and subunit movements compared to the initial structure are 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 employed an iterative combination of MDFF with Rosetta\textsuperscript{126,127}. Here, we used the optimized strategy as described for model construction of the 26S proteasome\textsuperscript{128,129}.

To obtain an atomic model that fit 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 significant 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 significant 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.

Last, the PSII-M, PSII-I, and PSII-I’ models were used to initiate one final round of real-space refinement in Phenix\textsuperscript{130}.

Acknowledgements

We thank C. König, M. Vökel, and R. Oworah-Nkruma for excellent technical assistance, Kristin Becker for cloning of the pIVEXPsb28His plasmid, Bibi Erjavec for preparation of the scheme in Fig. 1 and Nicholas 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. **Funding:** Financial support was provided by the Max Planck Society, the Helmholtz Zentrum München, the DFG research unit FOR2092 (EN 1194/1-1 to B.D.E., NO 836/3-2 to M.M.N.), the DFG priority program 2002 (NO 836/4-1 to M.M.N.), the grant 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 number ANR-10-LABX-0040-SPS) and the French Infrastructure for Integrated Structural Biology (FRISBI; grant number ANR-10-INSB-05). R.S. gratefully acknowledges support from the DFG (INST 213/757-1 FUGG and INST 213/843-1 FUGG). **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 contributions of 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 and 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 analyzed by O.A. and R.S.. All authors approved the final version of the manuscript. **Competing interests:** The authors declare no competing interests. **Data availability:** The cryo-EM density maps will be deposited in the Electron Microscopy Data Bank, the atomic models of the cryo-EM structures in the worldwide Protein Data Bank (wwPDB) and the NMR assignments for Psb28 in the Biological Magnetic Resonance Bank (BMRB), respectively.
References


52 Johnson, G. N., Boussac, A. & Rutherford, A. W. The origin of 40–50 C
thermoluminescence bands in photosystem II. Biochim Biophys Acta 1184, 85-92
(1994).
53 Stowell, M. H. et al. Light-induced structural changes in photosynthetic reaction center:
implications for mechanism of electron-proton transfer. Science 276, 812-816,
54 Cormann, K. U., Möller, M. & Nowaczyk, M. M. Critical Assessment of Protein Cross-
Linking and Molecular Docking: An Updated Model for the Interaction Between
55 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 U S A 108, 18536-18541,
56 Wei, X. et al. Structure of spinach photosystem II-LHCII supercomplex at 3.2 A
57 Cormann, K. U. et al. Structure of Psb27 in solution: implications for transient binding
to photosystem II during biogenesis and repair. Biochemistry 48, 8768-8770,
58 Fagerlund, R. D. & Eaton-Rye, J. J. The lipoproteins of cyanobacterial photosystem II.
59 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,
60 Michoux, F. et al. Crystal structure of the Psb27 assembly factor at 1.6 A: implications
for binding to Photosystem II. Photosynth Res 110, 169-175, doi:10.1007/s11120-011-
9712-7 (2012).
to Dissect the Life-Cycle of Photosystem II. Front Plant Sci 7, 617,
62 Kettunen, R., Tyystjarvi, E. & Aro, E. M. 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,
63 Mulo, P., Laakso, S., Maenpaa, P. & Aro, E. M. 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,
64 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,
65 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,
66 Bao, H. & Burnap, R. L. Photoactivation: The Light-Driven Assembly of the Water
Oxidation Complex of Photosystem II. Front Plant Sci 7, 578,
67 Becker, K., Cormann, K. U. & Nowaczyk, M. M. Assembly of the water-oxidizing
complex in photosystem II. J Photochem Photobiol B 104, 204-211,


