



**HAL**  
open science

## 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, et al.

► **To cite this version:**

Jure Zabret, Stefan Bohn, Sandra K Schuller, Oliver Arnolds, Madeline Möller, et al.. Structural insights into photosystem II assembly. *Nature Plants*, 2021, 7 (4), pp.524-538. 10.1038/s41477-021-00895-0 . hal-03406233

**HAL Id: hal-03406233**

**<https://hal.science/hal-03406233v1>**

Submitted on 27 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Structural insights into photosystem II assembly

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

### Affiliations:

<sup>1</sup>Department of Plant Biochemistry, Faculty of Biology & Biotechnology, Ruhr University Bochum, 44780 Bochum, Germany.

<sup>2</sup>Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany.

<sup>3</sup>Department of Structural Cell Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany.

<sup>4</sup>CryoEM of Molecular Machines, SYNMIKRO Research Center and Department of Chemistry, Philipps University of Marburg, 35043 Marburg, Germany.

<sup>5</sup>Biomolecular Spectroscopy and RUBiospek|NMR, Faculty of Chemistry and Biochemistry, Ruhr-University Bochum, 44780 Bochum, Germany

<sup>6</sup>Department of Molecular Membrane Biology, Max Planck Institute of Biophysics, 60438 Frankfurt/Main, Germany.

<sup>7</sup>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, Illinois

<sup>8</sup>Max Planck Institute for Brain Research, Max von Laue Strasse 4, 60438 Frankfurt/Main, Germany.

<sup>9</sup>Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

<sup>10</sup>Helmholtz Pioneer Campus, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany.

<sup>11</sup>Department of Chemistry, Technical University of Munich, Lichtenbergstraße 4, 85748 Garching, Germany.

<sup>12</sup>Biospectroscopy, Center for Protein Diagnostics (ProDi), Ruhr University Bochum, 44801 Bochum, Germany.

<sup>13</sup>Department of Biophysics, Faculty of Biology & Biotechnology, Ruhr University Bochum, 44780 Bochum, Germany.

#Equal contribution

\*corresponding authors: [till.rudack@rub.de](mailto:till.rudack@rub.de), [jan.schuller@synmikro.uni-marburg.de](mailto:jan.schuller@synmikro.uni-marburg.de), [marc.m.nowaczyk@rub.de](mailto:marc.m.nowaczyk@rub.de)

## Keywords

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

## 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, 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<sub>B</sub>) and replace the bicarbonate ligand of non-heme iron with glutamate, 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<sub>4</sub>CaO<sub>5</sub> cluster, which performs the unique water splitting reaction.

## 1 Introduction

2 Photosystem II (PSII) is the only enzyme that catalyzes the light-driven oxidation of water, a  
3 thermodynamically demanding reaction that drives photosynthesis, sustaining life on our  
4 planet<sup>1-3</sup>. This multi-subunit membrane protein complex is located in the thylakoid membranes  
5 of cyanobacteria, algae and plants. PSII strips electrons from water and injects them into the  
6 photosynthetic electron transport chain (PET). It forms a homodimer with a molecular mass of  
7 ~500 kDa<sup>4</sup>, with each monomer composed of at least 20 protein subunits and numerous  
8 cofactors, including chlorophylls, quinones, carotenoids, lipids, bicarbonate and the unique  
9 Mn<sub>4</sub>CaO<sub>5</sub> cluster<sup>5-7</sup>. The two core proteins D1 and D2 form a central, membrane-intrinsic  
10 heterodimer, which binds all important redox cofactors involved in internal electron transfer<sup>8</sup>.  
11 Light-excitation leads to a charge-separated state in which an electron is transferred from the  
12 reaction center chlorophylls P<sub>680</sub><sup>9</sup> to the nearby pheophytin<sup>10</sup>. Subsequently, the electron is  
13 passed to the bound plastoquinone (Q<sub>A</sub>) and then to the mobile plastoquinone molecule (Q<sub>B</sub>),  
14 which leaves the complex after accepting two electrons and two protons<sup>11</sup>. The electron hole at  
15 P<sub>680</sub> is filled by oxidation of an adjacent tyrosine residue (Tyr<sub>Z</sub>)<sup>12</sup> and finally by the oxygen

16 evolving complex (OEC), which contains the  $Mn_4CaO_5$  cluster. In cyanobacteria, the cluster is  
17 shielded on the luminal side by the three extrinsic proteins, PsbO, PsbU and PsbV, which  
18 regulate access to the OEC by forming an intricate network of channels for different substrates  
19 and products<sup>13</sup>. Light energy is collected and funneled towards  $P_{680}$  by the two membrane-  
20 intrinsic antenna proteins CP43 and CP47. These proteins bind most of the chlorophyll  
21 molecules and are located on opposite sides of the D1/D2 heterodimer<sup>14</sup>. Moreover, at least  
22 twelve small transmembrane subunits with one or two transmembrane helices have been  
23 identified in PSII<sup>15</sup>, including cytochrome- $b_{559}$ <sup>16</sup>.

24 Structural and spectroscopic investigations have revealed these aforementioned comprehensive  
25 insights into PSII function<sup>17-21</sup>, but we are far from understanding PSII biogenesis with  
26 molecular detail. How nature facilitates the assembly of a multi-subunit, multi-cofactor  
27 membrane protein complex is a fundamental unsolved question. The biogenesis of PSII is even  
28 more challenging, as the mature complex performs sophisticated and extreme redox chemistry  
29 to catalyze the light-driven oxidation of water. This can easily lead to the formation of reactive  
30 oxygen species (e.g., singlet oxygen is produced by triplet chlorophyll in the PSII reaction  
31 center) and subsequent loss of function due to damaged proteins and cofactors<sup>22,23</sup>. Biogenesis  
32 intermediates with only partially functional fragments of the redox chain are particularly prone  
33 to damage, thus demanding specialized protection mechanisms for the assembly process<sup>24</sup>.  
34 Therefore, PSII biogenesis is not a spontaneous process but rather must be tightly regulated by  
35 the action of assembly factors. Thus far, more than 20 auxiliary proteins have been identified  
36 that guide the stepwise assembly of PSII subunits and cofactors via intermediate modules,  
37 which are assembled independently and then joined together to produce mature PSII<sup>25-27</sup>. In  
38 cyanobacteria, PSII biogenesis begins with the formation of the D1/D2 heterodimer reaction  
39 center (RC) complex from the D1 precursor protein (pD1) and the D2 protein. This is assisted  
40 by the PSII assembly factor Ycf48 after partial processing of the D1 C-terminal extension by  
41 the D1 specific peptidase CtpA<sup>28,29</sup>. In the next step, the assembly factor Psb28 helps CP47 join  
42 the RC complex to form the RC47 complex, in which iD1 is further processed to its mature  
43 form by CtpA<sup>30,31</sup>. Almost all ligands of the  $Mn_4CaO_5$  cluster are already present at this stage,  
44 except for those provided by CP43, which comes pre-constructed with assembly factor Psb27  
45 and several small subunits (together called the CP43 module)<sup>32</sup>. Psb28 is released as CP43  
46 binds, and the resulting Psb27-PSII monomer is activated by maturation of the OEC and the  
47 binding of the extrinsic proteins PsbO, PsbU and PsbV<sup>33-35</sup>. OEC assembly is a multistep  
48 process that requires a functional upstream redox chain for the oxidation of  $Mn^{2+}$  to build up  
49 the cluster's  $\mu$ -oxo bridges between the manganese atoms<sup>36,37</sup>. The mechanistic and structural



50 details of this photoactivation process are not yet understood. In the consensus ‘two quantum  
51 model’<sup>38-40</sup>, a single Mn<sup>2+</sup> ion bound to the high-affinity site (HAS)<sup>41</sup> is oxidized to Mn<sup>3+</sup>. This  
52 initiating light-dependent step is followed by a slow light-independent phase and further fast  
53 light-dependent steps in which the remaining Mn<sup>2+</sup> ions are oxidized and incorporated.  
54 Understanding the light-independent slow phase is key to unraveling the mechanism of  
55 photoactivation. Finally, PSII biogenesis completes with dimerization of two fully assembled  
56 monomers and attachment of the soluble phycobilisome antenna complexes. Interestingly,  
57 deletion of *psbJ*, which encodes a small single transmembrane helix protein at the entrance of  
58 the PSII plastoquinone channel, decreases PSII-mediated oxygen evolution and increases the  
59 lifetime of the reduced primary acceptor (Q<sub>A</sub><sup>-</sup>) in cyanobacterial and tobacco chloroplasts<sup>42</sup>.  
60 These effects are caused by a massive accumulation of an intermediate monomeric PSII  
61 complex, which contains both assembly factors Psb27 and Psb28<sup>43</sup>. Physiological studies of  
62 Psb27 and Psb28 deletion strains point towards multifaceted functions. Cyanobacterial mutants  
63 lacking Psb28 exhibited slower autotrophic growth, particularly under stress conditions<sup>31,44</sup>, and  
64 limited synthesis of Chl-binding proteins but without decreased PSII functionality<sup>31</sup>. The Psb28  
65 mutant also exhibited an overall increase in PSII repair and faster recovery from  
66 photodamage<sup>31</sup>. Chemical cross-linking combined with mass spectrometry revealed that Psb28  
67 binds to the cytosolic side of CP47 close to cytochrome-b<sub>559</sub> and the Q<sub>B</sub> binding site. Based on  
68 this, researchers postulated a protective role for Psb28, where it blocks electron transport to the  
69 acceptor side of PSII to shield the RC47 complex from excess photodamage during the  
70 assembly process<sup>45</sup>. This hypothesis is strengthened by the observation that Psb28 is also found  
71 in PSII repair complexes<sup>46</sup>. The luminal PSII assembly factor Psb27 has been similarly well  
72 investigated. This lipoprotein is predominantly associated with inactive PSII fractions involved  
73 in assembly or repair<sup>32,34,46-50</sup>, stabilizing the CP43 luminal domain and presumably facilitating  
74 the assembly of the OEC.

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

82 Here, we use cryo-EM single particle analysis to describe the first molecular structure of a PSII  
83 assembly intermediate. This structure represents one of the key transitions in PSII biogenesis:

84 the attachment of the CP43 module to the pre-assembled RC47 reaction center complex, which  
85 precedes incorporation and activation of the  $Mn_4CaO_5$  cluster. We complement this structural  
86 data with spectroscopic analysis, revealing the first detailed insights into the molecular  
87 mechanisms of PSII assembly. Our study provides mechanistic answers to three long-standing  
88 questions: i) How do assembly factors modulate the structures of PSII subunits to assist  
89 biogenesis? ii) How is PSII protected from photodamage during assembly? iii) How is the PSII  
90 active site prepared for incorporation of the  $Mn_4CaO_5$  cluster?

## 91 **Results**

92 Stable PSII intermediates were purified from the *T. elongatus*  $\Delta psbJ$  mutant<sup>43</sup> by affinity  
93 chromatography using a twin-strep-tag fused to the C-terminus of the CP43 subunit and  
94 subsequent ion exchange chromatography (Fig. S1a). The main peak of the IEC profile  
95 corresponds primarily to monomeric PSII, which lacks the extrinsic subunits PsbO, PsbU and  
96 PsbV that are indicative for water splitting activity (Fig. S1b and c). Single particle cryo-EM  
97 analysis of this PSII fraction resulted in three different high-resolution maps that allowed model  
98 building with high confidence and excellent statistics (Fig. S2, Table S1). In addition to the  
99 protein subunits, we also faithfully assigned all essential non-protein cofactors, including  
100 chlorophylls, quinones, carotenoids and lipids, which are also present in the mature PSII  
101 complex (Fig. S3). Consistent with previous biochemical studies<sup>34,43,48</sup>, the EM density  
102 corresponding to the fully assembled, active  $Mn_4CaO_5$  cluster is missing in the purified  
103 biogenesis intermediates. The first cryo-EM map (2.94 Å), which we call PSII-I (for PSII-  
104 Intermediate), provides a snapshot of the attachment of the CP43 module to the pre-assembled  
105 RC47 reaction center complex (Fig. 1). This PSII intermediate contains three assembly factors  
106 (Psb27, Psb28 and Psb34), as well as almost all the membrane-intrinsic subunits and cofactors  
107 found in mature PSII, although CP43 is bound in an immature conformation. Psb27 and Psb28  
108 are well-known assembly factors<sup>31,32,35,43,51</sup>, whereas the additional single transmembrane helix  
109 protein (tsl0063), which we named Psb34, has not been described before. The small subunit  
110 PsbY, which is known to be loosely bound<sup>52</sup>, is not resolved in our structure. In addition, PsbJ  
111 is not present, as the corresponding gene was inactivated to stall PSII assembly at this specific  
112 transition<sup>43</sup>.

113 The two additional maps serve as internal controls. PSII-I' (2.76 Å) lacks Psb27 but is otherwise  
114 comparable to PSII-I; the root mean square deviation (RMSD) of the  $C_\alpha$  atomic positions  
115 between similar subunits of the two complexes is 0.4 Å. Most likely, Psb27 was partly lost

116 during sample preparation. The third cryo-EM map (2.82 Å), which we call PSII-M (for PSII-  
117 Monomer), represents a monomeric PSII complex without bound assembly factors.  
118 Comparison of our PSII-M structure with a crystal structure of monomeric PSII<sup>52</sup> (PDB-ID  
119 3KZI, 3.6 Å) reveals only minimal differences between both structures, with a C $\alpha$  RMSD of  
120 1.3 Å, which verifies that the structural changes observed in PSII-I are not caused by the  
121 deletion of PsbJ.

### 122 **Psb34 binds to CP47 and is involved in the RC47 transition**

123 Our PSII-I structure provides the first identification of the single transmembrane helix protein  
124 Psb34 bound to a PSII assembly intermediate (Fig. 2a), which we also confirmed by mass  
125 spectrometry (Fig. 2b). Psb34 was probably overlooked previously due to its hydrophobicity  
126 and small size. It has a single transmembrane helix that binds to the CP47 antenna protein in  
127 close proximity to PsbH (Fig. 2a). Its conserved long N-terminal arm is located at the side and  
128 top of the D2 subunit (Fig. 2a). In addition, we independently confirmed the interaction of  
129 Psb34 with PSII assembly intermediates by isolation of strep-tagged Psb34 complexes from  
130 otherwise unmodified cells, indicating a specific function of Psb34 in the attachment of CP43  
131 to RC47 (Fig. 2c). In this case, two distinct PSII intermediates were isolated via pulldown of  
132 strep-tagged Psb34: the RC47 complex with bound Psb28 and the subsequent PSII intermediate  
133 after attachment of CP43 and Psb27 (Fig. 2c). This observation implies that Psb28 is usually  
134 released from the PSII intermediate after attachment of CP43, probably after incorporation of  
135 PsbJ, as this trigger is missing in the analyzed  $\Delta psbJ$  mutant. PSII-I, which only accumulates  
136 in the  $\Delta psbJ$  mutant, seems to represent a hybrid state of both PSII intermediates that might be  
137 stabilized by the incomplete binding of CP43. Psb34 shows sequence similarity to high-light  
138 inducible proteins (HLIPs), which play a role in transient chlorophyll storage and chlorophyll  
139 biosynthesis<sup>53</sup>. However, the chlorophyll binding motive is missing in Psb34 (Table S2),  
140 suggesting a distinct function for this protein in PSII biogenesis.

### 141 **Psb28 induces conformational changes to the D1 D-E loop**

142 Psb28 binds on the cytosolic faces of the D1 and D2 subunits, directly above the Q<sub>B</sub> binding  
143 site (Fig. 3a), which differs from the position that was previously predicted by mass  
144 spectrometry<sup>45</sup>. Its binding induces the formation of an extended beta-hairpin structure that  
145 incorporates the central anti-parallel beta-sheet of Psb28, the C-terminus of CP47 and the D1  
146 D-E loop<sup>54</sup> (Fig. 3a). Binding of Psb28 to the C-terminus of CP47 also imparts a directionality

147 to the assembly process. In the Psb28-free complex (PSII-M), the CP47 C-terminus blocks the  
148 Psb28 binding site by interacting with the D1 D-E loop, thus preventing the reverse process and  
149 perturbation of active PSII by Psb28. Using nuclear magnetic resonance (NMR) spectroscopy,  
150 we performed chemical shift perturbation (CSP) experiments with recombinant Psb28 and a  
151 synthetic peptide of the conserved CP47 C-terminus to characterize this interaction in detail  
152 and determine the dissociation constant ( $K_D$ ) (Fig. 3 and Fig. S4). The CSP measurements  
153 indicated significant shifts with a chemical shift difference ( $\Delta\delta$ ) of more than one standard  
154 deviation located at strands  $\beta_3$  and  $\beta_4$  as well at the C-terminal region of Psb28 (Fig. 3c and  
155 d). Upon peptide binding, resonances for several residues gradually appeared with increasing  
156 peptide concentration, which were line-broadened beyond detection for the free form of Psb28.  
157 This observation indicates a less dynamic and more rigid complex structure. This is further  
158 supported by the heteronuclear Overhauser effect (NOE) data, which show that the C-terminus  
159 of Psb28 becomes rigid from L108 to K112 upon CP47 peptide binding due to creation of an  
160 intermolecular  $\beta$ -sheet (Fig. 3e). 2D-lineshape analysis was performed, yielding a  $K_D$  of  
161  $53.92 \pm 0.41 \mu\text{M}$  and a dissociation rate  $k_{\text{off}}$  of  $10.14 \pm 0.16 \text{ s}^{-1}$ , which is consistent with the  
162 observed slow-exchange in the NMR spectra (Fig. 3b). The affinity of Psb28 for full-length  
163 CP47 and PSII might indeed be even higher due to additional contacts between Psb28 and the  
164 D-E loop of D1 (Fig 3a).

### 165 **Psb28 prevents full association of CP43 and distorts the $Q_B$ -site**

166 Binding of Psb28—with support of Psb34—causes major structural perturbations at the PSII  
167 acceptor side (Supplementary Movies 1 and 2), which mainly involve the D-E loops of the  
168 central D1 and D2 subunits. Comparison of the CP43 structure in PSII-I with that in our  
169 PsbJ-free control PSII-M (Fig. 4a-d) or with that in mature monomeric PSII (PDB-ID 3KZI)  
170 (Fig. 4c and d) reveals several differences. The CP43 C-terminus is not resolved in PSII-I,  
171 probably due to an immature position of the last transmembrane helix of CP43 and an altered  
172 conformation of the D1 D-E loop, which may prevent binding of the CP43 C-terminus to the  
173 cytoplasmic PSII surface (Fig. 4b). This region is close to the loop between helices D and E of  
174 the D2 subunit, which is also altered by binding of Psb28, as clearly shown by movement of  
175 D2 Arg233 (Fig. 4b, Fig. S5a and b). After dissociation of Psb28, the CP43 module undergoes  
176 a rigid body rotation where it clicks into place (Fig. 4b-d, Supplementary Movie 1), whereas  
177 binding of PsbJ and the extrinsic proteins PsbO, PsbV and PsbU during further maturation has  
178 very little influence on the CP43 binding position (Fig. 4c and d). The part of PSII that

179 originates from RC47 shows almost no difference between PSII-I and mature PSII (Fig. 4d),  
180 except for PsbE, which binds adjacent to PsbJ (Fig. 4c).

181 Most importantly, the structural changes in the D1 D-E loop may have a direct functional impact  
182 on PSII electron transfer (Fig. 4e and f), as this region coordinates several important PSII  
183 cofactors. In functional PSII, after charge separation at the reaction center P<sub>680</sub>, electrons are  
184 transferred via pheophytin to the bound plastoquinone (Q<sub>A</sub>) and further to mobile plastoquinone  
185 (Q<sub>B</sub>). In our PSII-I structure, the Q<sub>A</sub> site is fully assembled, and a well-resolved Q<sub>A</sub> molecule is  
186 bound (Fig. 4e and f, Fig. S5c and d). The nearby non-heme iron is also already in place in  
187 PSII-I (Fig. 4e and f, Fig. S5e and f). The Q<sub>B</sub> binding site of the PSII-M control is comparable  
188 to mature PSII, although it is not occupied by Q<sub>B</sub> in our preparation (Fig. S5g). In contrast, the  
189 Q<sub>B</sub> binding site of PSII-I is immature due to the Psb28- and Psb34-induced structural changes  
190 in the D1 D-E loop (Fig. 4e and f, Fig. S5h). Notably, D1 Phe265, which coordinates the head  
191 group of Q<sub>B</sub> in mature PSII, is clearly at a different position<sup>21</sup> (Supplementary Movie 2).

## 192 **Binding of Psb28 protects PSII during biogenesis**

193 A more detailed analysis of the structural environment close to the Q<sub>A</sub>/Q<sub>B</sub> binding sites revealed  
194 differences in the coordination and the hydrogen-bond network of the adjacent non-heme iron,  
195 which also indicate functional consequences for PSII electron transfer and charge  
196 recombination processes. In mature PSII, the non-heme iron is coordinated by four histidine  
197 residues and bicarbonate as the fifth ligand (Fig. 5a and c), whereas in PSII-I, the bicarbonate  
198 molecule is replaced by the E241 side-chain of D2 (Fig. 5b and d, Fig. S5e and f, Supplementary  
199 Movie 3). Other residues, including D1 E244 and Y246, which bind to the bicarbonate molecule  
200 in mature PSII (Fig. 5a), are also displaced in PSII-I due to the conformational change of the  
201 D1 D-E loop (Fig. 5b, Fig. S5e and f, Supplementary Movie 3). Binding of bicarbonate is  
202 important for PSII efficiency<sup>55</sup>, as it lowers the redox potential of (Q<sub>A</sub>/Q<sub>A</sub><sup>-</sup>) to favor forward  
203 electron transport<sup>56,57</sup>. If charge recombination occurs, the lower redox potential favors indirect  
204 charge recombination via P<sub>680</sub><sup>+</sup>/Pheo<sup>-</sup>. This back reaction yields triplet chlorophyll and  
205 subsequently singlet oxygen<sup>57</sup>, a highly reactive oxidizing species. Changes in the redox  
206 potential of (Q<sub>A</sub>/Q<sub>A</sub><sup>-</sup>) have been proposed to tune the efficiency of PSII depending on the  
207 availability of CO<sub>2</sub> as the final electron acceptor and thereby protect PSII under low CO<sub>2</sub>  
208 conditions<sup>57</sup>. Therefore, we used flash-induced variable fluorescence to monitor the stability  
209 of Q<sub>A</sub><sup>-</sup> in active PSII, in the PSII-I assembly intermediate and in inactivated, Mn-depleted PSII,  
210 both of which lack a functional OEC (Fig. 5e, Fig. S6a and b). The results indicate that the

211 unique fluorescence traces observed with PSII-I are a consequence of  $Q_A/Q_A^-$  redox tuning by  
212 an absent OEC, an immature  $Q_B$ -binding pocket and probably the replacement of bicarbonate  
213 by glutamate as a ligand of the non-heme iron (see supplementary discussion for details). To  
214 determine the functional consequences of this redox tuning in PSII-I, we measured the  
215 formation of  $^1O_2$  by EPR spectroscopy using the spin probe TEMPD. The data clearly show  
216 that  $^1O_2$  formation is reduced by ~30% in PSII-I compared to inactivated PSII (Fig. 5f).

217

### 218 **Psb27 binds to loop E of CP43 at the luminal PSII surface**

219 Psb27 binds to the luminal side of the PSII complex, adjacent to loop E of the CP43 subunit  
220 (Fig. 6a and b). We could successfully model the Psb27 N-terminus until Cys22, which is the  
221 first residue after N-terminal processing of Psb27<sup>34</sup>. Psb27-Cys22 is located close to the  
222 membrane plane, but the lipid moiety is not well enough resolved to build a model. In contrast  
223 to previously proposed models<sup>58,59</sup>, the binding site of Psb27 has little overlap with the binding  
224 sites of the extrinsic subunits (PsbO, PsbV and PsbU) and has at least no direct contact to the  
225  $Mn_4CaO_5$  cluster binding site, particularly not to the D1 C-terminus (Fig. 6a and b), which has  
226 been suggested previously<sup>58,60</sup>. Instead, Psb27 is bound at a remote position that might be  
227 occupied by CyanoQ in the mature complex<sup>61</sup>. This position of Psb27 may still prevent proper  
228 binding of PsbO<sup>34</sup> by lowering its binding affinity due to the partial overlap of the binding sites,  
229 although both proteins can bind together, as PsbO-containing Psb27-PSII complexes have been  
230 isolated via his-tagged Psb27<sup>49</sup>. Additionally, Psb27 might stabilize loop E of CP43 in the  
231 unassembled state and facilitate its binding to the D1 subunit. This is of particular importance,  
232 as loop E of CP43 provides Arg345 and Glu342 (Arg357 and Glu354 in previous publications,  
233 see figure caption), to position the  $Mn_4CaO_5$  cluster in mature PSII (Fig. 6b, dashed box).  
234 Glu342 serves a ligand, whereas Arg345 stabilizes the  $Mn_4CaO_5$  cluster by coordinating  
235 Asp170 of D1 through a hydrogen bond. Moreover, in the Psb27-bound state (PSII-I), the D1  
236 C-terminus, which is directly involved in coordination of the  $Mn_4CaO_5$  cluster<sup>21</sup>, is bound away  
237 from the cluster (Fig. 6c, Fig. S7), which might be an additional consequence of Psb27-induced  
238 stabilization of the CP43 E-loop. Thus, our PSII-I structure reveals not only how the Psb27  
239 protein binds to CP43 and thus stabilizes it, but also indicates a role for the maturation of the  
240 oxygen evolving cluster that is consistent with functional data from previous studies<sup>32,35,48</sup>.

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

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

253 The D1 C-terminus is an essential component for formation of the of OEC, as it provides several  
254 charged residues that are responsible for coordination of the chloride ion and the  $Mn_4CaO_5$   
255 cluster (Fig. 7a, b and d). The density for these C-terminal residues is weak in our PSII-I map,  
256 but traceable (Fig. S7a), indicating a flexibility that confirms the absence of the OEC.  
257 Compared to the mature complex, the last 12 residues of the C-terminal tail of D1 would need  
258 to undergo significant conformational changes to bring the side chains of Glu333, His337,  
259 Asp342, and the Ala344 C-terminus into the correct position to coordinate the  $Mn_4CaO_5$  cluster  
260 (Supplementary Movie 4).

261 Moreover, we identify a clearly visible density at the position of the chloride ion, which is  
262 coordinated by Lys317 (D2) and the hydrogen atom of the backbone nitrogen of Glu333 (D1)  
263 in mature PSII (Fig. 7b and e). Despite the similar position, the  $Cl^-$  is coordinated by the nitrogen  
264 atom of the ring of adjacent His332 (D1) in PSII-I (Fig. 7a and e, Fig. S7d). Surprisingly, we  
265 identified another density in the area where the  $Mn_4CaO_5$  cluster is located in mature PSII  
266 (Fig. 7a-c and f, Fig. S7c). However, this density is not large enough to reflect the whole cluster.  
267 Based on its size and interaction partners (Fig. 7f), it corresponds to one positively charged ion.  
268 In the structural context, this ion is most likely  $Mn^{2+}$ , but it could also be  $Ca^{2+}$  or any other  
269 positively charged ion.

270 This ion is coordinated by the side chains of D1 Asp170, Glu189, and His332, which are already  
271 in similar positions compared to mature PSII. Glu342 and Arg345 (Glu354 and Arg357 in  
272 previous publications) of CP43, which are both involved in the second coordination sphere of  
273 the  $Mn_4CaO_5$  cluster, are also already pre-positioned through the interaction between Arg345

274 with D1 Asp170 (Fig. 7g). However, there are still significant conformational changes  
275 necessary for the transition from PSII-I to mature PSII, as highlighted in Figure 7d and g, as  
276 well as in Supplementary Movie 4. The D1 C-terminal tail must bring the side chains of Glu333,  
277 His337 and Asp342, as well as the C-terminus of Ala344, into correct alignment to coordinate  
278 the  $Mn_4CaO_5$  cluster. In addition, the C-terminal tail of D2 needs to flip towards the D1 C-  
279 terminus (Fig. 7c, Fig. S7b, Supplementary Movie 4). In summary, PSII-I is characterized by  
280 only one positive charged ion bound instead of the complete  $Mn_4CaO_5$  cluster, resulting in  
281 significantly different conformations of the D1 and D2 C-termini compared to the structural  
282 model containing a mature  $Mn_4CaO_5$  cluster. However, the PSII-I structure seems to be  
283 prepared to accept the  $Mn_4CaO_5$  cluster, as indicated by the above described similarities in side  
284 chain positioning.

## 285 **Discussion**

286 PSII biogenesis is a complex process that requires the action of specific assembly factors. These  
287 auxiliary proteins are not present in the mature complex and interact only transiently with  
288 specific subunits or preassembled PSII intermediates. Although more than 20 factors have been  
289 identified and assigned to specific transitions, their precise molecular functions in PSII  
290 assembly remain elusive in almost all cases. Our study provides the first detailed molecular  
291 insights into the function of PSII assembly factors Psb27, Psb28 and Psb34, which are involved  
292 in an important transition prior to activation of the OEC. The determined binding positions of  
293 Psb27 and Psb28, which are two of the most studied PSII assembly factors, define the binding  
294 with atomic resolution, in contrast to previous models<sup>58,59,62-66</sup>. This provides precise  
295 information about the molecular function of Psb27 and Psb28 in the assembly process.

296 Binding of Psb28 and Psb34 to the cytoplasmic side of PSII induces large conformational  
297 changes in the D1 D-E loop (Fig. 4), which has been identified previously as an important  
298 location for PSII photoinhibition and D1 degradation<sup>67,68</sup>. Structural changes observed in the  
299 PSII-I  $Q_B$  binding pocket and coordination of the non-heme iron suggest a functional impact on  
300 PSII electron transfer to protect the immature complex until water splitting is activated. In  
301 particular, D2 Glu241 replacing bicarbonate as a ligand of non-heme iron suggests a regulatory  
302 role, as binding of bicarbonate was proposed to tune PSII efficiency by changing the redox  
303 potential of ( $Q_A/Q_A^-$ )<sup>56,57</sup>. As a functional consequence, PSII-I generates less singlet oxygen  
304 compared to inactivated PSII (Fig. 5f).



305 Interestingly, the coordination of the non-heme iron in PSII-I resembles that in non-oxygenic  
306 bacterial reaction centers (BRCs)<sup>69</sup> (Fig. S6c). In BRCs, the fifth ligand of the non-heme iron  
307 is provided by E234 of the M subunit<sup>70</sup>, and mutagenesis of this residue induces changes in the  
308 free energy gap between the P<sup>•+</sup>/Q<sub>A</sub><sup>•-</sup> radical pair<sup>71</sup>. These findings indicate that the environment  
309 of the non-heme iron is important for the Q<sub>A</sub> redox potential, which influences charge  
310 recombination and the formation of reactive oxygen species<sup>56</sup>.

311 The Psb27-bound and -unbound structures do not differ substantially (Fig. 6b), suggesting a  
312 rather subtle action in PSII biogenesis. Previous work demonstrated that Psb27 is already bound  
313 to free CP43<sup>32</sup>, where it might protect free CP43 from degradation or stabilize the E-loop in a  
314 specific conformation to chaperone the subsequent association with the RC47 complex. This  
315 step is crucial for preparing the binding site of the Mn<sub>4</sub>CaO<sub>5</sub> cluster, as the CP43 E-loop is  
316 involved in its second coordination sphere<sup>72</sup>. Moreover, a recent high-speed AFM study with  
317 spinach PSII has shown fluctuations of the C43 E-loop at room temperature after removal of  
318 the extrinsic subunits and destruction of the Mn<sub>4</sub>CaO<sub>5</sub> cluster<sup>73</sup>, which might suggest a role of  
319 Psb27 in restraining the scope of conformational fluctuations for further assembly of the OEC.  
320 Consistent with this model, recent spectroscopic analysis of PSII from *Synechocystis* sp. PCC  
321 6803 revealed that Psb27 facilitates photoactivation, probably by stabilization of intermediates  
322 with increased accessibility for ions, especially Ca<sup>2+</sup>, to the site of cluster assembly<sup>74</sup>. However,  
323 our data indicate a self-organization of the CP43 E-loop, at least in *T. elongatus*. The  
324 corresponding structural models for this region are almost identical, independent of binding of  
325 Psb27 or the extrinsic proteins, whereas the positions of the D1 and D2 C-termini are clearly  
326 different between intermediate and mature PSII. Binding of PsbO seems to be the main driver  
327 for these structural differences, and its presence interferes with efficient photoactivation,  
328 probably by closing the cluster binding site and thus preventing the exchange of ions<sup>74</sup>. The  
329 overlap of the Psb27 and PsbO binding sites, although small, might decrease the PsbO affinity  
330 and prevent its premature binding, as initially suggested<sup>34,37</sup>, or Psb27 may promote binding of  
331 PsbO in a different conformation (both proteins can bind simultaneously in a stable complex<sup>49</sup>)  
332 that might support photoactivation.

333 Previous structural studies aimed to obtain mechanistic insights into the dark-rearrangement  
334 during photoactivation by removing the Mn<sub>4</sub>CaO<sub>5</sub> cluster from fully assembled PSII, either by  
335 depleting it directly from PSII crystals by chemical treatment<sup>75</sup> or by cryo-EM single particle  
336 analysis in manganese- and calcium-free buffer<sup>76</sup>. The X-ray structure was indeed missing the  
337 Mn<sub>4</sub>CaO<sub>5</sub> cluster, but the D1 C-terminus followed mostly the same trajectory as found in the  
338 mature PSII-dimer structure. The authors suggested that the D1 C-terminus might not rearrange

339 during  $Mn_4CaO_5$  cluster assembly. However, the crystal structure was dimeric and still had the  
340 extrinsic subunits PsbO, PsbU, and PsbV bound. It is known that these subunits are typically  
341 not associated with  $Mn_4CaO_5$  cluster-depleted PSII. Thus, the structure might be artificially  
342 stabilized by crystal packing forces. The cryo-EM structure, on the other hand, revealed a  
343 monomeric PSII that lacks extrinsic subunits and the  $Mn_4CaO_5$  cluster<sup>76</sup>. This structure is more  
344 similar to our PSII biogenesis intermediate PSII-I, as PsbY, PsbZ and PsbJ are also missing.  
345 The PsbJ subunit is surprising; it is an integral subunit of PSII and should not be easily detached,  
346 yet it is missing from this structure, as we deleted it to stabilize our PSII-I complex. These  
347 observations might indicate a more specific and regulatory role of PsbJ in PSII biogenesis.  
348 Additionally, the D1 C-terminus is disordered in this previous cryo-EM structure, and the  
349 authors suggest that the dark-rearrangement involves a transition from a disordered to an  
350 ordered state.

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

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

## 382 **Methods**

### 383 **Cultivation of *Thermosynechococcus elongatus* BP-1**

384 Cell growth and thylakoid membrane preparation were performed as described previously<sup>83</sup>. In  
385 brief, *T. elongatus* mutant strains ( $\Delta$ psbJ psbC-TS and psb34-TS) were grown in BG-11 liquid  
386 medium inside a 25-litre foil fermenter (Bioengineering) at 45°C, 5% (v/v) CO<sub>2</sub>-enriched air  
387 bubbling and 50-200  $\mu$ mol photons m<sup>-2</sup> s<sup>-1</sup> white light illumination (depending on the cell  
388 density). Cells were harvested at an OD<sub>680</sub> of ~ 2 after 5-6 days of cultivation and concentrated  
389 to ~ 0.5 l, using an Amicon DC10 LA hollow fibre system, pelleted (3500 rcf, 45 min and 25  
390 °C) and resuspended in 150 ml of Buffer D (100 mM Tris-HCL, pH 7.5, 10 mM MgCl<sub>2</sub>, 10 mM  
391 CaCl<sub>2</sub>, 500 mM mannitol and 20% (w/v) glycerol). The harvested cells were flash-frozen in  
392 liquid nitrogen and stored at -80 °C until further use.

### 393 **Preparation of *T. elongatus* mutant strains**

394 *Thermosynechococcus elongatus*  $\Delta$ psbJ psbC-TS was generated based on the previously  
395 described strain *T. elongatus*  $\Delta$ psbJ<sup>43</sup> that was transformed with the plasmid pCP43-TS. The  
396 plasmid is based on pCP34-10His<sup>34</sup>. The His-tag sequence was exchanged with TwinStrep-tag  
397 by PCR using the primers CP43TS\_rev (5'CCCGATATCTTACTTCTCAAATTGCGGAT  
398 GAGACCACGCAGAACCACCAAGACCGCCGCTGCCGCCGCCTTTTTTCGAACTG  
399 CGGGTGGCTCC 3') and NTCP43 (5' TGCTCTAGAATGAAAACCTTTGTCTTCCCAGA  
400 3'). The resulting PCR product was ligated back into an empty pCP34-10His backbone using  
401 XbaI and EcoRV restriction endonucleases. *T. elongatus* BP-1 cells were transformed as  
402 described previously<sup>84</sup>. Mutant colonies were selected by frequent re-plating onto agar plates  
403 with increasing antibiotic concentrations, stopping at 8  $\mu$ g/ml of chloramphenicol and 80  $\mu$ g/ml

404 of kanamycin. Complete segregation of the mutant was confirmed by PCR with the primers  
405 CTC43DS (5' CCGCTCGAGACCATCCAAGCTTGGCAGCA 3') and NTCP43 (5'  
406 TGCTCTAGAATGAAAACCTTTGTCTTCCCAGA 3').

407 *T. elongatus* psb34-TS was generated by transformation with the plasmid pPsb34-TS. The  
408 plasmid DNA was obtained from TwistBioscience. It consisted of psb34 (tsl0063) with a C-  
409 terminal TwinStrep-tag and a kanamycin resistance cassette, flanked by tsl0063-upstream and  
410 downstream regions (900 bp each). *T. elongatus* BP-1 cells were transformed<sup>84</sup> and mutant  
411 selection took place<sup>34</sup>. Complete segregation of the mutant was verified by PCR. The primers  
412 used were tsl0063-up-for (5' CATATGGTCTCGCAATTATTTGCCCATGC 3') and tsl0063-  
413 down-rev (5' GGTACCCCG ACACAGTTGATCACCGC 3').

#### 414 **Purification of photosystem II assembly intermediates**

415 Thawed cells were diluted in 100 ml of Buffer A (100 mM Tris-HCL, pH 7.5, 10 mM MgCl<sub>2</sub>  
416 and 10 mM CaCl<sub>2</sub>) and pelleted again (21 000 rcf, 20 min and 4°C). The pellet was resuspended  
417 in 100 ml of Buffer A with 0.2% (w/v) lysozyme and dark incubated for 75-90 min at 37 °C.  
418 This was followed by cell disruption by Parr bomb (Parr Instruments Company) and pelleting  
419 (21 000 rcf, 20 min and 4°C). All following steps were performed under green illumination.  
420 The pellet was resuspended in 150 ml of Buffer A and pelleted again (21 000 rcf, 20 min and  
421 4°C). This step was repeated three times, with the last resuspension in 80 ml of Buffer B (100  
422 mM Tris-HCL, pH 7.5, 10 mM MgCl<sub>2</sub>, 10 mM CaCl<sub>2</sub> and 500 mM mannitol). The isolated  
423 thylakoid membranes were flash frozen in liquid nitrogen and stored at -80 °C.

424 Strep-Tactin-affinity purification of PsbC-TS and Psb34-TS assembly intermediates were  
425 performed under green illumination. Membrane protein extraction was performed as described  
426 previously<sup>83</sup>, with certain adaptations. Isolated thylakoid membranes were supplemented with  
427 0.05% (w/v) n-Dodecyl β-maltoside (DDM) (Glycon) and pelleted (21000 rcf, 20 min and 4°C).  
428 The sample was resuspended in extraction buffer (100 mM Tris-HCL, pH 7.5, 10 mM MgCl<sub>2</sub>,  
429 10 mM CaCl<sub>2</sub>, 1.2% (w/v) DDM, 0.5% (w/v) sodium-cholate and 0.01% (w/v) DNase) to a  
430 final chlorophyll concentration of 1 mg/ml and incubated for 30 min at 20 °C. The solubilized  
431 membrane proteins were ultra-centrifugated (140000 rcf, 60 min and 4 °C) and NaCl was added  
432 to the supernatant to a final concentration of 300 mM.

433 The supernatant was filtered through a 0.45 μm syringe filter (Sarsted AG, Germany) and  
434 applied to a 5 ml Strep-Tactin Superflow HC column (IBA Lifesciences), equilibrated in Buffer  
435 W (100 mM Tris-HCL, pH 7.5, 10 mM MgCl<sub>2</sub>, 10 mM CaCl<sub>2</sub>, 500 mM mannitol, 300 mM

436 NaCl and 0.03% (w/v) DDM) at a flowrate of 3 ml/min. The column was washed with Buffer  
437 W until a stable baseline ( $A_{280}$ ) was reached. Strep-tagged protein complexes were eluted by an  
438 isocratic elution with Buffer E (100 mM Tris-HCL, pH 7.5, 10 mM  $MgCl_2$ , 10 mM  $CaCl_2$ , 500  
439 mM mannitol, 300 mM NaCl 2.5 mM desthiobiotin and 0.03% (w/v) DDM). The captured  
440 fractions were equilibrated in Buffer F (20 mM MES, pH 6.5, 10 mM  $MgCl_2$ , 10 mM  $CaCl_2$ ,  
441 500 mM mannitol and 0.03% (w/v) DDM) with a spin concentrator (Amicon, Ultra – 15,  
442 100000 NMWL), flash-frozen in liquid nitrogen and stored at  $-80\text{ }^\circ\text{C}$  until analysis.

443 PsbC-TS containing assembly intermediates were further separated by ion exchange  
444 chromatography (IEC). Captured elution fraction from the Strep-Tactin-affinity purification  
445 were loaded onto a anion exchange column (UNO Q-6, Biorad) with a flowrate of 4 ml/min,  
446 pre-equilibrated in Buffer F. Protein complexes were eluted by a liner gradient of  $MgSO_4$  (0-  
447 150 mM) using Buffer G (20 mM MES, pH 6.5, 10 mM  $MgCl_2$ , 10 mM  $CaCl_2$ , 500 mM  
448 mannitol, 150 mM  $MgSO_4$  and 0.03% (w/v) DDM). Fractions containing PSII assembly  
449 intermediates were collected, concentrated to 100 – 10  $\mu\text{M}$  reaction centers, using a spin  
450 concentrator (Amicon, Ultra – 15, 100 000 NMWL), aliquoted, flash frozen in liquid nitrogen  
451 and stored at  $-80\text{ }^\circ\text{C}$  until further analysis.

#### 452 **Cryo-electron microscopy**

453 For cryo-EM sample preparation, 4.5  $\mu\text{l}$  of purified protein complexes were applied to glow  
454 discharged Quantifoil 2/1 grids, blotted for 3.5 s with force 4 in a Vitrobot Mark III (Thermo  
455 Fisher) at 100% humidity and  $4\text{ }^\circ\text{C}$ , then plunge frozen in liquid ethane, cooled by liquid  
456 nitrogen. Cryo-EM data was acquired with a FEI Titan Krios transmission electron microscope  
457 using the SerialEM software<sup>85</sup>. Movie frames were recorded at a nominal magnification of  
458 22,500x using a K3 direct electron detector (Gatan), The total electron dose of  $\sim 55$  electrons  
459 per  $\text{\AA}^2$  was distributed over 30 frames at a calibrated physical pixel size of 1.09  $\text{\AA}$ . Micrographs  
460 were recorded in a defocus range of  $-0.5$  to  $-3.0\text{ }\mu\text{m}$ .

#### 461 **Image processing, classification and refinement**

462 Cryo-EM micrographs were processed on the fly using the Focus software package<sup>86</sup> if they  
463 passed the selection criteria (iciness  $< 1.05$ , drift  $0.4\text{ }\text{\AA} < x < 70\text{ }\text{\AA}$ , defocus  $0.5\text{ }\mu\text{m} < x < 5.5$   
464  $\mu\text{m}$ , estimated CTF resolution  $< 6\text{ }\text{\AA}$ ). Micrograph frames were aligned using MotionCorr<sup>87</sup> and  
465 the contrast transfer function (CTF) for aligned frames was determined using Gctf<sup>88</sup>. Using  
466 Gautomatch (<http://www.mrc-lmb.cam.ac.uk/kzhang/>) 693,297 particles were picked template-  
467 free on 824 acquired micrographs. Particles were extracted with a pixel box size of 260 using  
468 RELION 3.1<sup>89</sup> and imported into Cryosparc 2.3<sup>90</sup>. After reference-free 2D classification,

469 675,123 particles were used for ab initio construction of initial models and subjected to multiple  
470 rounds of 3D classification to obtain models with and without Psb28 density. Non-uniform  
471 refinement in Cryosparc resulted in models with an estimated resolution of  $\sim 3.2$  Å. Particles  
472 belonging to 3D classes with and without Psb28 (150,090 and 166,411 particles, respectively)  
473 were reextracted in RELION with a pixel box size of 256 and subjected to several rounds of  
474 CTF-refinement (estimation of anisotropic magnification, fit of per-micrograph defocus and  
475 astigmatism and beam tilt estimation) and Bayesian polishing<sup>91</sup>. Both classes were refined using  
476 the previously generated starting models. 3D classification without further alignment using a  
477 mask around the Psb27 region separated particles in the Psb28-containing class into distinct  
478 classes with and without Psb27 (57,862 and 91,473 particles, respectively). Final refinement of  
479 each of the three classes (with Psb27 and Psb28 (PSII-I), with Psb28 but without Psb27 (PSII-  
480 I'), and without Psb27 and Psb28 (PSII-M)) resulted in models with global resolutions of 2.94  
481 Å, 2.76 Å and 2.82 Å, respectively (Gold standard FSC analysis of two independent half-sets  
482 at the 0.143 cutoff). Local-resolution and 3D-FSC plots (Extended Data Fig. 2) were calculated  
483 using RELION and the “Remote 3DFSC Processing Server” web interface<sup>92</sup>, respectively.

#### 484 **Atomic model construction**

485 The 3.6 Å resolution X-ray structure of monomeric PSII from *T. elongatus* with PDB-ID 3KZI<sup>52</sup>  
486 was used as initial structural model that was docked as rigid body using Chimera<sup>93</sup> into the  
487 obtained cryo EM densities for PSII-M and PSII-I. The cofactors that had no corresponding  
488 density were removed. The subunit PsbJ was also removed, as it was deleted in the experimental  
489 design. By highlighting the still unoccupied parts of the PSII-I density map, we identified  
490 densities that lead to the structures of Psb27, Psb28, and Psb34.

491 The 2.4 Å resolution X-ray structures of isolated Psb28 from *T. elongatus* with PDB-ID 3ZPN<sup>94</sup>  
492 and the 1.6 Å resolution X-ray structure of isolated Psb27 from *T. elongatus* with PDB-ID  
493 2Y6X<sup>65</sup> were docked as rigid bodies into the unoccupied densities. The 1.6 Å resolution X-ray  
494 structure of CyanoQ from *T. elongatus* with PDB-ID 3ZSU<sup>95</sup> does not fit into the density and  
495 was therefore not modeled.

496 As there was no experimentally resolved structural model of Psb34 available, we first used the  
497 sequence with UniProt-ID Q8DMP8 to predict structures using the webserver SWISS Model<sup>96</sup>  
498 and LOMETS<sup>97</sup>. We also predicted the secondary structure through the meta server  
499 Bioinformatics Toolkit<sup>98</sup> and CCTOP<sup>99</sup>. The results of the secondary structure prediction are  
500 summarized in Table S4. Combining these predictions together with the unassigned cryo-EM

501 density, we used COOT<sup>100</sup> to build an initial model of Psb34 that has one  $\alpha$ -helix from amino  
502 acid number 28 to 55.

### 503 **Model Refinement**

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

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

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

528 Last, the PSII-M, PSII-I, and PSII-I' models were used to initiate one final round of real-space  
529 refinement in Phenix<sup>109</sup>.

530

### 531 **Acknowledgements**

532 We thank C. König, M. Völkel, and R. Oworah-Nkruma for excellent technical assistance,  
533 Kristin Becker for cloning of the pIVEX<sub>Psb28His</sub> plasmid, Bibi Erjavec for preparation of the

534 scheme in Fig. 1 and Nicholas Cox for helpful discussion. J.M.S. is grateful to E. Conti for  
535 scientific independence and great mentorship and to J. M. Plitzko and W. Baumeister for access  
536 to the cryo-EM infrastructure and early career support. M.M.N. is grateful to his mentor M.  
537 Rögner for generous support. **Funding:** Financial support was provided by the Max Planck  
538 Society, the Helmholtz Zentrum München, the DFG research unit FOR2092 (EN 1194/1-1 to  
539 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  
540 grant NIH P41-GM104601 (to E.T.) and an Emmy-Noether fellowship (SCHU 3364/1-1 to  
541 J.M.S). A.K.-L. was supported by the LabEx Saclay Plant Sciences-SPS (grant number ANR-  
542 10-LABX-0040-SPS) and the French Infrastructure for Integrated Structural Biology (FRISBI;  
543 grant number ANR-10-INSB-05). R.S. gratefully acknowledges support from the DFG (INST  
544 213/757-1 FUGG and INST 213/843-1 FUGG). **Author contributions:** B.D.E., T.R., J.M.S.  
545 and M.M.N. conceived the research, prepared the figures, and wrote the manuscript with the  
546 contributions of all other authors. M.M.N. coordinated the activities. Preparation of mutants,  
547 PSII isolation and biochemical analysis were performed by J.Z., M.M., P.L. and M.M.N. Mass  
548 spectrometry analysis was done by J.M.-C. and J.D.L. J.M.S., S.B. and B.D.E. performed the  
549 cryo-EM analysis and T.R. built the structural model with the help of S.K.S., A.C. and E.T.  
550 Fluorescence spectroscopy was carried out by J.Z. and M.M.N. EPR experiments were  
551 conducted by A.K.-L. NMR experiments were conducted and analyzed by O.A. and R.S.. All  
552 authors approved the final version of the manuscript. **Competing interests:** The authors declare  
553 no competing interests. **Data availability:** The cryo-EM density maps will be deposited in the  
554 Electron Microscopy Data Bank, the atomic models of the cryo-EM structures in the worldwide  
555 Protein Data Bank (wwPDB) and the NMR assignments for Psb28 in the Biological Magnetic  
556 Resonance Bank (BMRB), respectively.

## 557 **References**

- 558 1 Hohmann-Marriott, M. F. & Blankenship, R. E. Evolution of photosynthesis. *Annual*  
559 *review of plant biology* **62**, 515-548, doi:10.1146/annurev-arplant-042110-103811  
560 (2011).
- 561 2 Sanchez-Baracaldo, P. & Cardona, T. On the origin of oxygenic photosynthesis and  
562 Cyanobacteria. *New Phytol* **225**, 1440-1446, doi:10.1111/nph.16249 (2020).
- 563 3 Vinyard, D. J., Ananyev, G. M. & Dismukes, G. C. Photosystem II: the reaction center  
564 of oxygenic photosynthesis. *Annual review of biochemistry* **82**, 577-606,  
565 doi:10.1146/annurev-biochem-070511-100425 (2013).
- 566 4 Boekema, E. J. *et al.* Supramolecular structure of the photosystem II complex from  
567 green plants and cyanobacteria. *Proc Natl Acad Sci U S A* **92**, 175-179,  
568 doi:10.1073/pnas.92.1.175 (1995).



- 569 5 Cox, N., Pantazis, D. A. & Lubitz, W. Current Understanding of the Mechanism of  
570 Water Oxidation in Photosystem II and Its Relation to XFEL Data. *Annual review of*  
571 *biochemistry* **89**, 795-820, doi:10.1146/annurev-biochem-011520-104801 (2020).
- 572 6 Shen, J. R. The Structure of Photosystem II and the Mechanism of Water Oxidation in  
573 Photosynthesis. *Annual review of plant biology* **66**, 23-48, doi:10.1146/annurev-arplant-  
574 050312-120129 (2015).
- 575 7 Yano, J. *et al.* Light-dependent production of dioxygen in photosynthesis. *Met Ions Life*  
576 *Sci* **15**, 13-43, doi:10.1007/978-3-319-12415-5\_2 (2015).
- 577 8 Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J. & Iwata, S. Architecture of  
578 the photosynthetic oxygen-evolving center. *Science* **303**, 1831-1838,  
579 doi:10.1126/science.1093087 (2004).
- 580 9 Cardona, T., Sedoud, A., Cox, N. & Rutherford, A. W. Charge separation in  
581 photosystem II: a comparative and evolutionary overview. *Biochim Biophys Acta* **1817**,  
582 26-43, doi:10.1016/j.bbabi.2011.07.012 (2012).
- 583 10 Holzwarth, A. R. *et al.* Kinetics and mechanism of electron transfer in intact  
584 photosystem II and in the isolated reaction center: pheophytin is the primary electron  
585 acceptor. *Proc Natl Acad Sci U S A* **103**, 6895-6900, doi:10.1073/pnas.0505371103  
586 (2006).
- 587 11 Müh, F., Glöckner, C., Hellmich, J. & Zouni, A. Light-induced quinone reduction in  
588 photosystem II. *Biochim Biophys Acta* **1817**, 44-65, doi:10.1016/j.bbabi.2011.05.021  
589 (2012).
- 590 12 Faller, P. *et al.* Rapid formation of the stable tyrosyl radical in photosystem II. *Proc*  
591 *Natl Acad Sci U S A* **98**, 14368-14373, doi:10.1073/pnas.251382598 (2001).
- 592 13 Roose, J. L., Frankel, L. K., Mummadisetti, M. P. & Bricker, T. M. The extrinsic  
593 proteins of photosystem II: update. *Planta* **243**, 889-908, doi:10.1007/s00425-015-  
594 2462-6 (2016).
- 595 14 Müh, F. & Zouni, A. Structural basis of light-harvesting in the photosystem II core  
596 complex. *Protein Sci* **29**, 1090-1119, doi:10.1002/pro.3841 (2020).
- 597 15 Shi, L. X., Hall, M., Funk, C. & Schröder, W. P. Photosystem II, a growing complex:  
598 updates on newly discovered components and low molecular mass proteins. *Biochim*  
599 *Biophys Acta* **1817**, 13-25, doi:10.1016/j.bbabi.2011.08.008 (2012).
- 600 16 Stewart, D. H. & Brudvig, G. W. Cytochrome b559 of photosystem II. *Biochim Biophys*  
601 *Acta* **1367**, 63-87, doi:10.1016/s0005-2728(98)00139-x (1998).
- 602 17 Cox, N. *et al.* Photosynthesis. Electronic structure of the oxygen-evolving complex in  
603 photosystem II prior to O-O bond formation. *Science* **345**, 804-808,  
604 doi:10.1126/science.1254910 (2014).
- 605 18 Kern, J. *et al.* Structures of the intermediates of Kok's photosynthetic water oxidation  
606 clock. *Nature* **563**, 421-425, doi:10.1038/s41586-018-0681-2 (2018).
- 607 19 Kupitz, C. *et al.* Serial time-resolved crystallography of photosystem II using a  
608 femtosecond X-ray laser. *Nature* **513**, 261-265, doi:10.1038/nature13453 (2014).
- 609 20 Suga, M. *et al.* An oxyl/oxo mechanism for oxygen-oxygen coupling in PSII revealed  
610 by an x-ray free-electron laser. *Science* **366**, 334-338, doi:10.1126/science.aax6998  
611 (2019).
- 612 21 Umena, Y., Kawakami, K., Shen, J. R. & Kamiya, N. Crystal structure of oxygen-  
613 evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55-60,  
614 doi:10.1038/nature09913 (2011).
- 615 22 Krieger-Liszak, A., Fufezan, C. & Trebst, A. Singlet oxygen production in  
616 photosystem II and related protection mechanism. *Photosynth Res* **98**, 551-564,  
617 doi:10.1007/s11120-008-9349-3 (2008).
- 618 23 Vass, I. Molecular mechanisms of photodamage in the Photosystem II complex.  
619 *Biochim Biophys Acta* **1817**, 209-217, doi:10.1016/j.bbabi.2011.04.014 (2012).

- 620 24 Shevela, D. *et al.* ‘Birth defects’ of photosystem II make it highly susceptible to  
621 photodamage during chloroplast biogenesis. *Physiologia plantarum* **166**, 165-180  
622 (2019).
- 623 25 Heinz, S., Liauw, P., Nickelsen, J. & Nowaczyk, M. Analysis of photosystem II  
624 biogenesis in cyanobacteria. *Biochim Biophys Acta* **1857**, 274-287,  
625 doi:10.1016/j.bbabi.2015.11.007 (2016).
- 626 26 Nickelsen, J. & Rengstl, B. Photosystem II assembly: from cyanobacteria to plants.  
627 *Annual review of plant biology* **64**, 609-635, doi:10.1146/annurev-arplant-050312-  
628 120124 (2013).
- 629 27 Nixon, P. J., Michoux, F., Yu, J., Boehm, M. & Komenda, J. Recent advances in  
630 understanding the assembly and repair of photosystem II. *Ann Bot* **106**, 1-16,  
631 doi:10.1093/aob/mcq059 (2010).
- 632 28 Komenda, J. *et al.* Cleavage after residue Ala352 in the C-terminal extension is an early  
633 step in the maturation of the D1 subunit of Photosystem II in *Synechocystis* PCC 6803.  
634 *Biochim Biophys Acta* **1767**, 829-837, doi:10.1016/j.bbabi.2007.01.005 (2007).
- 635 29 Komenda, J. *et al.* The cyanobacterial homologue of HCF136/YCF48 is a component  
636 of an early photosystem II assembly complex and is important for both the efficient  
637 assembly and repair of photosystem II in *Synechocystis* sp. PCC 6803. *J Biol Chem* **283**,  
638 22390-22399, doi:10.1074/jbc.M801917200 (2008).
- 639 30 Boehm, M. *et al.* Subunit composition of CP43-less photosystem II complexes of  
640 *Synechocystis* sp. PCC 6803: implications for the assembly and repair of photosystem  
641 II. *Philos Trans R Soc Lond B Biol Sci* **367**, 3444-3454, doi:10.1098/rstb.2012.0066  
642 (2012).
- 643 31 Dobáková, M., Sobotka, R., Tichy, M. & Komenda, J. Psb28 protein is involved in the  
644 biogenesis of the photosystem II inner antenna CP47 (PsbB) in the cyanobacterium  
645 *Synechocystis* sp. PCC 6803. *Plant Physiol* **149**, 1076-1086,  
646 doi:10.1104/pp.108.130039 (2009).
- 647 32 Komenda, J. *et al.* The Psb27 assembly factor binds to the CP43 complex of  
648 photosystem II in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol* **158**,  
649 476-486, doi:10.1104/pp.111.184184 (2012).
- 650 33 Mamedov, F., Nowaczyk, M. M., Thapper, A., Rögner, M. & Styring, S. Functional  
651 characterization of monomeric photosystem II core preparations from  
652 *Thermosynechococcus elongatus* with or without the Psb27 protein. *Biochemistry* **46**,  
653 5542-5551, doi:10.1021/bi7000399 (2007).
- 654 34 Nowaczyk, M. M. *et al.* Psb27, a cyanobacterial lipoprotein, is involved in the repair  
655 cycle of photosystem II. *Plant Cell* **18**, 3121-3131, doi:10.1105/tpc.106.042671 (2006).
- 656 35 Roose, J. L. & Pakrasi, H. B. The Psb27 protein facilitates manganese cluster assembly  
657 in photosystem II. *J Biol Chem* **283**, 4044-4050, doi:10.1074/jbc.M708960200 (2008).
- 658 36 Bao, H. & Burnap, R. L. Photoactivation: The Light-Driven Assembly of the Water  
659 Oxidation Complex of Photosystem II. *Front Plant Sci* **7**, 578,  
660 doi:10.3389/fpls.2016.00578 (2016).
- 661 37 Becker, K., Cormann, K. U. & Nowaczyk, M. M. Assembly of the water-oxidizing  
662 complex in photosystem II. *J Photochem Photobiol B* **104**, 204-211,  
663 doi:10.1016/j.jphotobiol.2011.02.005 (2011).
- 664 38 Cheniae, G. & Martin, I. Photoactivation of the manganese catalyst of O<sub>2</sub> evolution. I-  
665 Biochemical and kinetic aspects. *Biochim Biophys Acta* **253**, 167-181,  
666 doi:10.1016/0005-2728(71)90242-8 (1971).
- 667 39 Dasgupta, J., Ananyev, G. M. & Dismukes, G. C. Photoassembly of the Water-  
668 Oxidizing Complex in Photosystem II. *Coord Chem Rev* **252**, 347-360,  
669 doi:10.1016/j.ccr.2007.08.022 (2008).

670 40 Radmer, R. & Cheniae, G. M. Photoactivation of the manganese catalyst of O<sub>2</sub>  
671 evolution. II. A two-quantum mechanism. *Biochimica et Biophysica Acta* **253**, 182-186,  
672 doi: 10.1016/0005-2728(71)90243-X (1971).

673 41 Nixon, P. J. & Diner, B. A. Aspartate 170 of the photosystem II reaction center  
674 polypeptide D1 is involved in the assembly of the oxygen-evolving manganese cluster.  
675 *Biochemistry* **31**, 942-948, doi:10.1021/bi00118a041 (1992).

676 42 Regel, R. E. *et al.* Deregulation of electron flow within photosystem II in the absence  
677 of the PsbJ protein. *Journal of Biological Chemistry* **276**, 41473-41478 (2001).

678 43 Nowaczyk, M. M. *et al.* Deletion of *psbJ* leads to accumulation of Psb27-Psb28  
679 photosystem II complexes in *Thermosynechococcus elongatus*. *Biochim Biophys Acta*  
680 **1817**, 1339-1345, doi:10.1016/j.bbabi.2012.02.017 (2012).

681 44 Sakata, S., Mizusawa, N., Kubota-Kawai, H., Sakurai, I. & Wada, H. Psb28 is involved  
682 in recovery of photosystem II at high temperature in *Synechocystis* sp. PCC 6803.  
683 *Biochim Biophys Acta* **1827**, 50-59, doi:10.1016/j.bbabi.2012.10.004 (2013).

684 45 Weisz, D. A. *et al.* Mass spectrometry-based cross-linking study shows that the Psb28  
685 protein binds to cytochrome b559 in Photosystem II. *Proc Natl Acad Sci U S A* **114**,  
686 2224-2229, doi:10.1073/pnas.1620360114 (2017).

687 46 Bečková, M. *et al.* Association of Psb28 and Psb27 Proteins with PSII-PSI  
688 Supercomplexes upon Exposure of *Synechocystis* sp. PCC 6803 to High Light. *Mol*  
689 *Plant* **10**, 62-72, doi:10.1016/j.molp.2016.08.001 (2017).

690 47 Bentley, F. K., Luo, H., Dilbeck, P., Burnap, R. L. & Eaton-Rye, J. J. Effects of  
691 inactivating psbM and psbT on photodamage and assembly of photosystem II in  
692 *Synechocystis* sp. PCC 6803. *Biochemistry* **47**, 11637-11646, doi:10.1021/bi800804h  
693 (2008).

694 48 Grasse, N. *et al.* Role of novel dimeric Photosystem II (PSII)-Psb27 protein complex in  
695 PSII repair. *J Biol Chem* **286**, 29548-29555, doi:10.1074/jbc.M111.238394 (2011).

696 49 Liu, H., Roose, J. L., Cameron, J. C. & Pakrasi, H. B. A genetically tagged Psb27 protein  
697 allows purification of two consecutive photosystem II (PSII) assembly intermediates in  
698 *Synechocystis* 6803, a cyanobacterium. *J Biol Chem* **286**, 24865-24871,  
699 doi:10.1074/jbc.M111.246231 (2011).

700 50 Weisz, D. A. *et al.* A novel chlorophyll protein complex in the repair cycle of  
701 photosystem II. *Proc Natl Acad Sci U S A* **116**, 21907-21913,  
702 doi:10.1073/pnas.1909644116 (2019).

703 51 Mabbitt, P. D., Wilbanks, S. M. & Eaton-Rye, J. J. Structure and function of the  
704 hydrophilic Photosystem II assembly proteins: Psb27, Psb28 and Ycf48. *Plant Physiol*  
705 *Biochem* **81**, 96-107, doi:10.1016/j.plaphy.2014.02.013 (2014).

706 52 Broser, M. *et al.* Crystal structure of monomeric photosystem II from  
707 *Thermosynechococcus elongatus* at 3.6-Å resolution. *J Biol Chem* **285**, 26255-26262,  
708 doi:10.1074/jbc.M110.127589 (2010).

709 53 Komenda, J. & Sobotka, R. Cyanobacterial high-light-inducible proteins--Protectors of  
710 chlorophyll-protein synthesis and assembly. *Biochim Biophys Acta* **1857**, 288-295,  
711 doi:10.1016/j.bbabi.2015.08.011 (2016).

712 54 Mulo, P. *et al.* Mutagenesis of the D-E loop of photosystem II reaction centre protein  
713 D1. Function and assembly of photosystem II. *Plant Mol Biol* **33**, 1059-1071,  
714 doi:10.1023/a:1005765305956 (1997).

715 55 Eaton-Rye, J. J. & Govindjee. Electron transfer through the quinone acceptor complex  
716 of photosystem II in bicarbonate-depleted spinach thylakoid membranes as a function  
717 of actinic flash number and frequency. *Biochimica et Biophysica Acta (BBA)-*  
718 *Bioenergetics* **935**, 237-247 (1988).

719 56 Allen, J. F. & Nield, J. Redox Tuning in Photosystem II. *Trends Plant Sci* **22**, 97-99,  
720 doi:10.1016/j.tplants.2016.11.011 (2017).

721 57 Brinkert, K., De Causmaecker, S., Krieger-Liszkay, A., Fantuzzi, A. & Rutherford, A.  
722 W. Bicarbonate-induced redox tuning in Photosystem II for regulation and protection.  
723 *Proc Natl Acad Sci U S A* **113**, 12144-12149, doi:10.1073/pnas.1608862113 (2016).

724 58 Cormann, K. U., Möller, M. & Nowaczyk, M. M. Critical Assessment of Protein Cross-  
725 Linking and Molecular Docking: An Updated Model for the Interaction Between  
726 Photosystem II and Psb27. *Front Plant Sci* **7**, 157, doi:10.3389/fpls.2016.00157 (2016).

727 59 Liu, H., Huang, R. Y., Chen, J., Gross, M. L. & Pakrasi, H. B. Psb27, a transiently  
728 associated protein, binds to the chlorophyll binding protein CP43 in photosystem II  
729 assembly intermediates. *Proc Natl Acad Sci U S A* **108**, 18536-18541,  
730 doi:10.1073/pnas.1111597108 (2011).

731 60 Wei, L. *et al.* LPA19, a Psb27 homolog in *Arabidopsis thaliana*, facilitates D1 protein  
732 precursor processing during PSII biogenesis. *Journal of Biological Chemistry* **285**,  
733 21391-21398 (2010).

734 61 Wei, X. *et al.* Structure of spinach photosystem II-LHCII supercomplex at 3.2 Å  
735 resolution. *Nature* **534**, 69-74, doi:10.1038/nature18020 (2016).

736 62 Cormann, K. U. *et al.* Structure of Psb27 in solution: implications for transient binding  
737 to photosystem II during biogenesis and repair. *Biochemistry* **48**, 8768-8770,  
738 doi:10.1021/bi9012726 (2009).

739 63 Fagerlund, R. D. & Eaton-Rye, J. J. The lipoproteins of cyanobacterial photosystem II.  
740 *J Photochem Photobiol B* **104**, 191-203, doi:10.1016/j.jphotobiol.2011.01.022 (2011).

741 64 Liu, H. *et al.* Mass spectrometry-based footprinting reveals structural dynamics of loop  
742 E of the chlorophyll-binding protein CP43 during photosystem II assembly in the  
743 cyanobacterium *Synechocystis* 6803. *J Biol Chem* **288**, 14212-14220,  
744 doi:10.1074/jbc.M113.467613 (2013).

745 65 Michoux, F. *et al.* Crystal structure of the Psb27 assembly factor at 1.6 Å: implications  
746 for binding to Photosystem II. *Photosynth Res* **110**, 169-175, doi:10.1007/s1120-011-  
747 9712-7 (2012).

748 66 Weisz, D. A., Gross, M. L. & Pakrasi, H. B. The Use of Advanced Mass Spectrometry  
749 to Dissect the Life-Cycle of Photosystem II. *Front Plant Sci* **7**, 617,  
750 doi:10.3389/fpls.2016.00617 (2016).

751 67 Kettunen, R., Tyystjarvi, E. & Aro, E. M. Degradation pattern of photosystem II  
752 reaction center protein D1 in intact leaves. The major photoinhibition-induced cleavage  
753 site in D1 polypeptide is located amino terminally of the DE loop. *Plant Physiol* **111**,  
754 1183-1190, doi:10.1104/pp.111.4.1183 (1996).

755 68 Mulo, P., Laakso, S., Maenpaa, P. & Aro, E. M. Stepwise photoinhibition of  
756 photosystem II. Studies with *Synechocystis* species PCC 6803 mutants with a modified  
757 D-E loop of the reaction center polypeptide D1. *Plant Physiol* **117**, 483-490,  
758 doi:10.1104/pp.117.2.483 (1998).

759 69 Stowell, M. H. *et al.* Light-induced structural changes in photosynthetic reaction center:  
760 implications for mechanism of electron-proton transfer. *Science* **276**, 812-816,  
761 doi:10.1126/science.276.5313.812 (1997).

762 70 Wang, X. *et al.* Is bicarbonate in Photosystem II the equivalent of the glutamate ligand  
763 to the iron atom in bacterial reaction centers? *Biochim Biophys Acta* **1100**, 1-8,  
764 doi:10.1016/0005-2728(92)90119-m (1992).

765 71 Cheap, H. *et al.* M234Glu is a component of the proton sponge in the reaction center  
766 from photosynthetic bacteria. *Biochim Biophys Acta* **1787**, 1505-1515,  
767 doi:10.1016/j.bbabi.2009.07.004 (2009).

768 72 Burnap, R. L. D1 protein processing and Mn cluster assembly in light of the emerging  
769 Photosystem II structure. *Physical Chemistry Chemical Physics* **6**, 4803-4809 (2004).

770 73 Tokano, T., Kato, Y., Sugiyama, S., Uchihashi, T. & Noguchi, T. Structural dynamics  
771 of a protein domain relevant to the water-oxidizing complex in photosystem II as

772 visualized by high-speed atomic force microscopy. *The Journal of Physical Chemistry*  
773 *B* **124**, 5847-5857 (2020).

774 74 Avramov, A. P., Hwang, H. J. & Burnap, R. L. The role of Ca<sup>2+</sup> and protein scaffolding  
775 in the formation of nature's water oxidizing complex. *Proceedings of the National*  
776 *Academy of Sciences* **117**, 28036-28045 (2020).

777 75 Zhang, M. *et al.* Structural insights into the light-driven auto-assembly process of the  
778 water-oxidizing Mn<sub>4</sub>CaO<sub>5</sub>-cluster in photosystem II. *Elife* **6**, doi:10.7554/eLife.26933  
779 (2017).

780 76 Gisriel, C. J. *et al.* Cryo-EM Structure of Monomeric Photosystem II from  
781 *Synechocystis* sp. PCC 6803 Lacking the Water-Oxidation Complex. *Joule* (2020).

782 77 Kolling, D. R., Cox, N., Ananyev, G. M., Pace, R. J. & Dismukes, G. C. What are the  
783 oxidation states of manganese required to catalyze photosynthetic water oxidation?  
784 *Biophys J* **103**, 313-322, doi:10.1016/j.bpj.2012.05.031 (2012).

785 78 Zaltsman, L., Ananyev, G. M., Bruntrager, E. & Dismukes, G. C. Quantitative kinetic  
786 model for photoassembly of the photosynthetic water oxidase from its inorganic  
787 constituents: requirements for manganese and calcium in the kinetically resolved steps.  
788 *Biochemistry* **36**, 8914-8922, doi:10.1021/bi970187f (1997).

789 79 Stengel, A. *et al.* Initial steps of photosystem II de novo assembly and preloading with  
790 manganese take place in biogenesis centers in *Synechocystis*. *The Plant Cell* **24**, 660-  
791 675 (2012).

792 80 Tyryshkin, A. M. *et al.* Spectroscopic evidence for Ca<sup>2+</sup> involvement in the assembly  
793 of the Mn<sub>4</sub>Ca cluster in the photosynthetic water-oxidizing complex. *Biochemistry* **45**,  
794 12876-12889, doi:10.1021/bi061495t (2006).

795 81 Campbell, K. A. *et al.* Dual-mode EPR detects the initial intermediate in photoassembly  
796 of the photosystem II Mn cluster: the influence of amino acid residue 170 of the D1  
797 polypeptide on Mn coordination. *Journal of the American Chemical Society* **122**, 3754-  
798 3761 (2000).

799 82 Cohen, R. O., Nixon, P. J. & Diner, B. A. Participation of the C-terminal region of the  
800 D1-polypeptide in the first steps in the assembly of the Mn<sub>4</sub>Ca cluster of photosystem  
801 II. *J Biol Chem* **282**, 7209-7218, doi:10.1074/jbc.M606255200 (2007).

802 83 Kuhl, H. *et al.* Towards structural determination of the water-splitting enzyme.  
803 Purification, crystallization, and preliminary crystallographic studies of photosystem II  
804 from a thermophilic cyanobacterium. *J Biol Chem* **275**, 20652-20659,  
805 doi:10.1074/jbc.M001321200 (2000).

806 84 Iwai, M., Katoh, H., Katayama, M. & Ikeuchi, M. Improved genetic transformation of  
807 the thermophilic cyanobacterium, *Thermosynechococcus elongatus* BP-1. *Plant Cell*  
808 *Physiol* **45**, 171-175, doi:10.1093/pcp/pch015 (2004).

809 85 Mastronarde, D. N. Automated electron microscope tomography using robust prediction  
810 of specimen movements. *Journal of structural biology* **152**, 36-51 (2005).

811 86 Biyani, N. *et al.* Focus: The interface between data collection and data processing in  
812 cryo-EM. *J Struct Biol* **198**, 124-133, doi:10.1016/j.jsb.2017.03.007 (2017).

813 87 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for  
814 improved cryo-electron microscopy. *Nat Methods* **14**, 331-332,  
815 doi:10.1038/nmeth.4193 (2017).

816 88 Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12,  
817 doi:10.1016/j.jsb.2015.11.003 (2016).

818 89 Scheres, S. H. A Bayesian view on cryo-EM structure determination. *J Mol Biol* **415**,  
819 406-418, doi:10.1016/j.jmb.2011.11.010 (2012).

820 90 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms  
821 for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296,  
822 doi:10.1038/nmeth.4169 (2017).

823 91 Scheres, S. H. Beam-induced motion correction for sub-megadalton cryo-EM particles.  
824 *Elife* **3**, e03665, doi:10.7554/eLife.03665 (2014).

825 92 Tan, Y. Z. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM  
826 through tilting. *Nat Methods* **14**, 793-796, doi:10.1038/nmeth.4347 (2017).

827 93 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research  
828 and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).

829 94 Bialek, W. *et al.* Crystal structure of the Psb28 accessory factor of  
830 *Thermosynechococcus elongatus* photosystem II at 2.3 Å. *Photosynth Res* **117**, 375-  
831 383, doi:10.1007/s11120-013-9939-6 (2013).

832 95 Michoux, F. *et al.* Crystal structure of CyanoQ from the thermophilic cyanobacterium  
833 *Thermosynechococcus elongatus* and detection in isolated photosystem II complexes.  
834 *Photosynth Res* **122**, 57-67, doi:10.1007/s11120-014-0010-z (2014).

835 96 Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. SWISS-MODEL: An automated  
836 protein homology-modeling server. *Nucleic Acids Res* **31**, 3381-3385,  
837 doi:10.1093/nar/gkg520 (2003).

838 97 Wu, S. & Zhang, Y. LOMETS: a local meta-threading-server for protein structure  
839 prediction. *Nucleic Acids Res* **35**, 3375-3382, doi:10.1093/nar/gkm251 (2007).

840 98 Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with  
841 a New HHpred Server at its Core. *J Mol Biol* **430**, 2237-2243,  
842 doi:10.1016/j.jmb.2017.12.007 (2018).

843 99 Dobson, L., Remenyi, I. & Tusnady, G. E. CCTOP: a Consensus Constrained TOPology  
844 prediction web server. *Nucleic Acids Res* **43**, W408-412, doi:10.1093/nar/gkv451  
845 (2015).

846 100 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of  
847 Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501,  
848 doi:10.1107/S0907444910007493 (2010).

849 101 Trabuco, L. G., Villa, E., Schreiner, E., Harrison, C. B. & Schulten, K. Molecular  
850 dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-  
851 ray crystallography. *Methods* **49**, 174-180, doi:10.1016/j.ymeth.2009.04.005 (2009).

852 102 Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol*  
853 *Graph* **14**, 33-38, doi:10.1016/0263-7855(96)00018-5 (1996).

854 103 Ribeiro, J. V. *et al.* QwikMD - Integrative Molecular Dynamics Toolkit for Novices and  
855 Experts. *Sci Rep* **6**, 26536, doi:10.1038/srep26536 (2016).

856 104 Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem* **26**,  
857 1781-1802, doi:10.1002/jcc.20289 (2005).

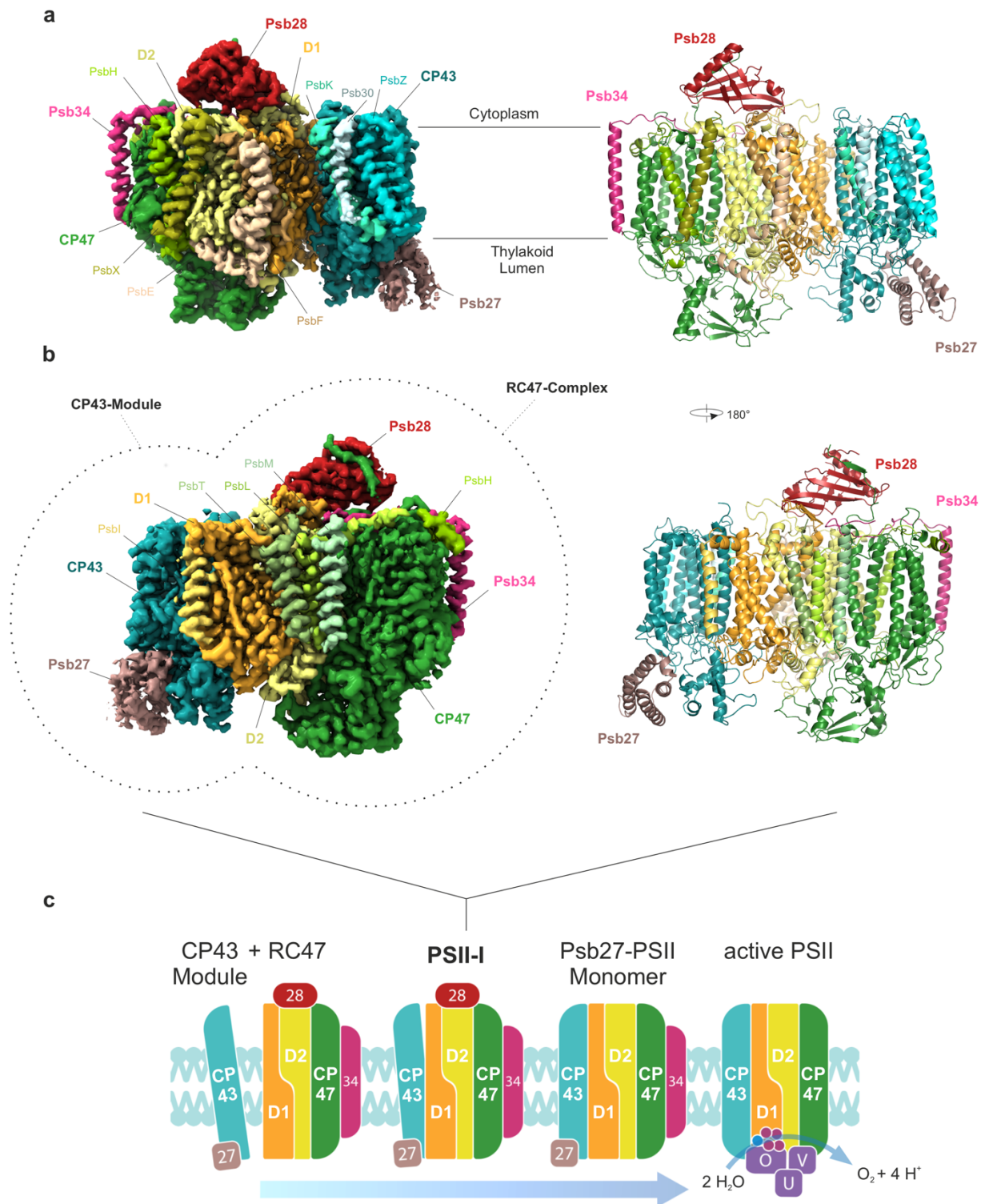
858 105 Leaver-Fay, A. *et al.* ROSETTA3: an object-oriented software suite for the simulation  
859 and design of macromolecules. *Methods Enzymol* **487**, 545-574, doi:10.1016/B978-0-  
860 12-381270-4.00019-6 (2011).

861 106 Lindert, S. & McCammon, J. A. Improved cryoEM-Guided Iterative Molecular  
862 Dynamics--Rosetta Protein Structure Refinement Protocol for High Precision Protein  
863 Structure Prediction. *J Chem Theory Comput* **11**, 1337-1346, doi:10.1021/ct500995d  
864 (2015).

865 107 Guo, Q. *et al.* In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals  
866 Proteasome Recruitment. *Cell* **172**, 696-705 e612, doi:10.1016/j.cell.2017.12.030  
867 (2018).

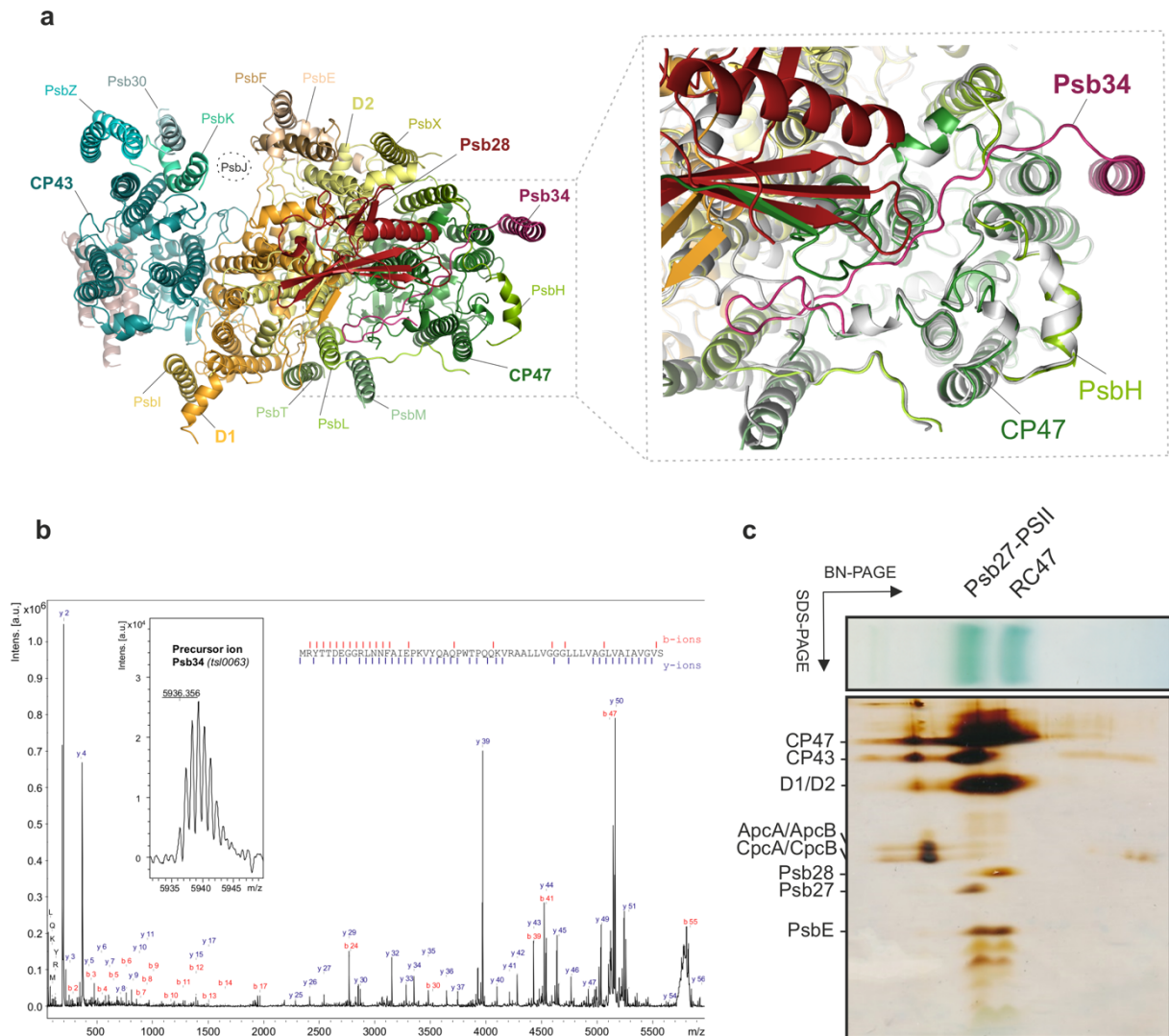
868 108 Wehmer, M. *et al.* Structural insights into the functional cycle of the ATPase module of  
869 the 26S proteasome. *Proc Natl Acad Sci U S A* **114**, 1305-1310,  
870 doi:10.1073/pnas.1621129114 (2017).

871 109 Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons  
872 and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-  
873 877, doi:10.1107/S2059798319011471 (2019).



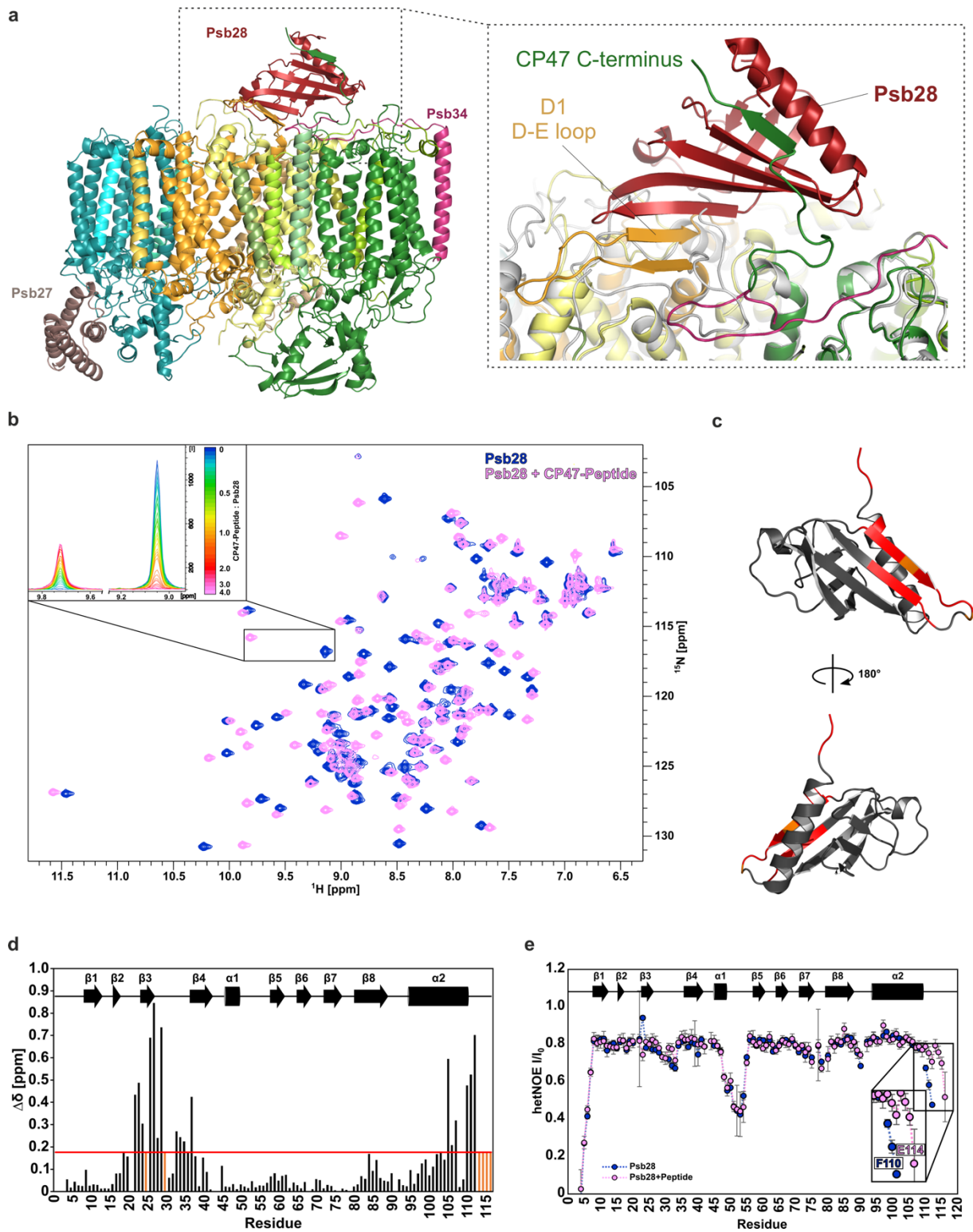
**Fig. 1: Cryo-EM map of a PSII assembly intermediate (PSII-I) from *T. elongatus*, segmented by subunit. a**, 15 PSII subunits and 3 assembly factors are colored and named (PSII subunits: D1, D2, CP43, CP47, PsbE, PsbF, PsbH, PsbI, PsbK, PsbL, PsbM, PsbT, PsbX, PsbZ and Psb30; assembly factors: Psb27, Psb28 and *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.





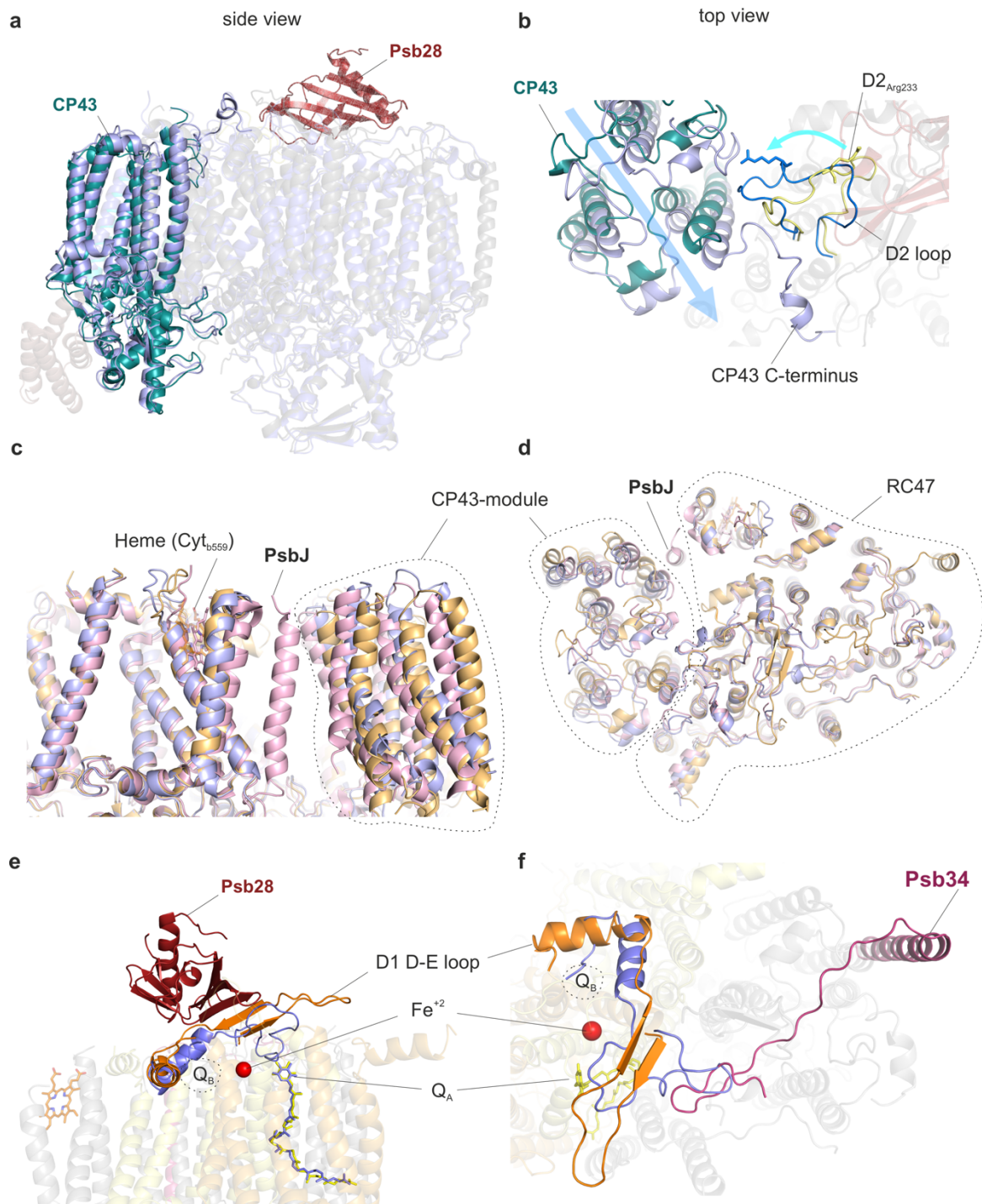
**Fig. 2: Psb34 binds to RC47 during attachment of the CP43 module. a**, Binding site of Psb34 at CP47, close to PsbH (top view), with extended binding of the Psb34 N-terminus along the cytoplasmic PSII surface (dashed box). **b**, MALDI-ToF analysis of PSII assembly intermediates. Mass spectrum of Psb34 (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.



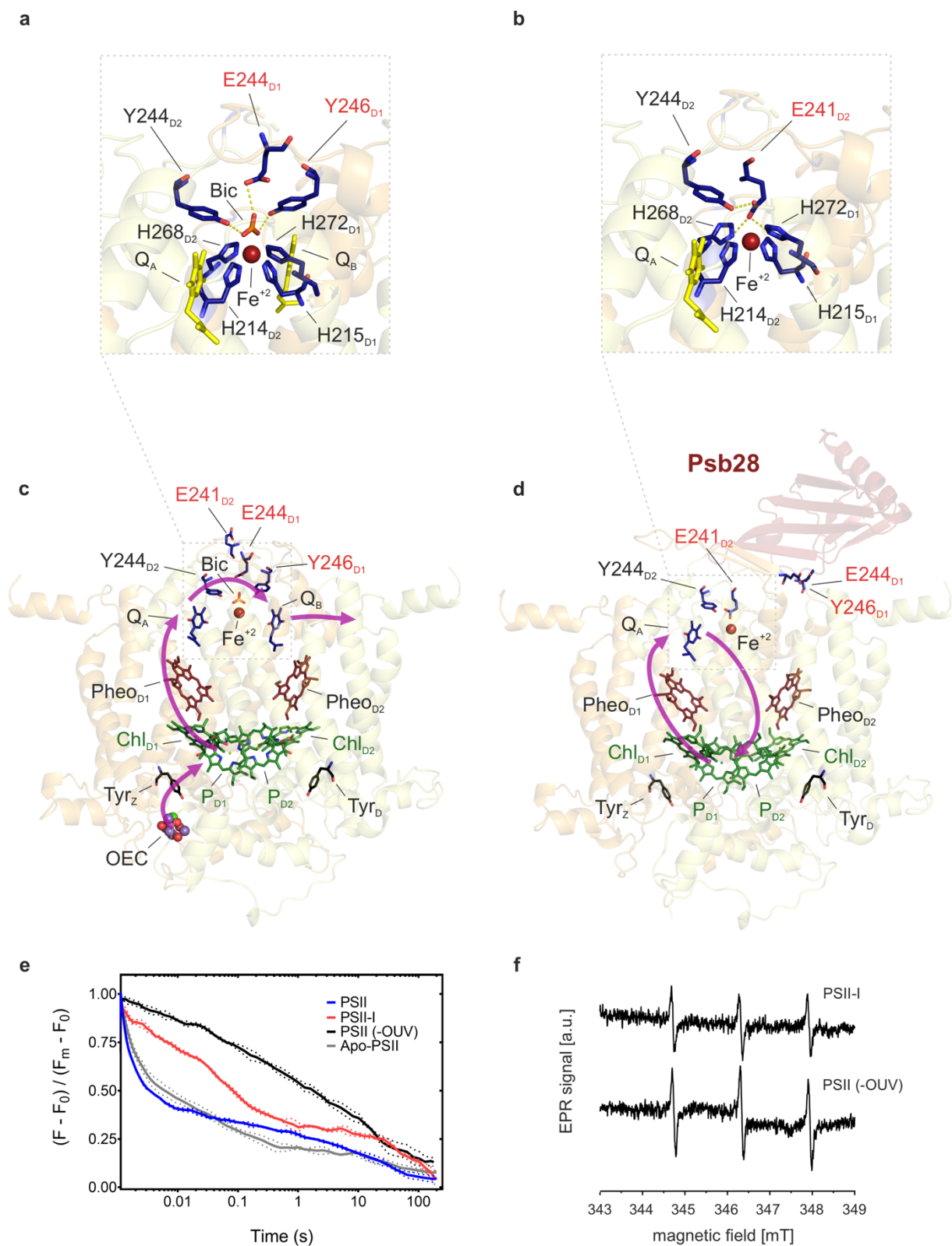


**Fig. 3: The role of the CP47 C-terminus in binding of Psb28.** **a**, Binding of Psb28 at the cytoplasmic/stromal PSII surface (side view, colors correspond to Fig. 1) and continuation of the central Psb28 beta-sheet by the CP47 C-terminus and the D-E loop of D1 (dashed box). For comparison, mature monomeric PSII (PDB-ID 3KZI) is shown in gray. **b**, Superimposed 2D  $^1\text{H}$ - $^{15}\text{N}$ -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  $^{15}\text{N}$  dimension. **c**, CSPs of more than one SD projected onto the model representation of Psb28. **d**, Weighted  $^1\text{H}/^{15}\text{N}$  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  $^{15}\text{N}$   $\{^1\text{H}\}$ -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.



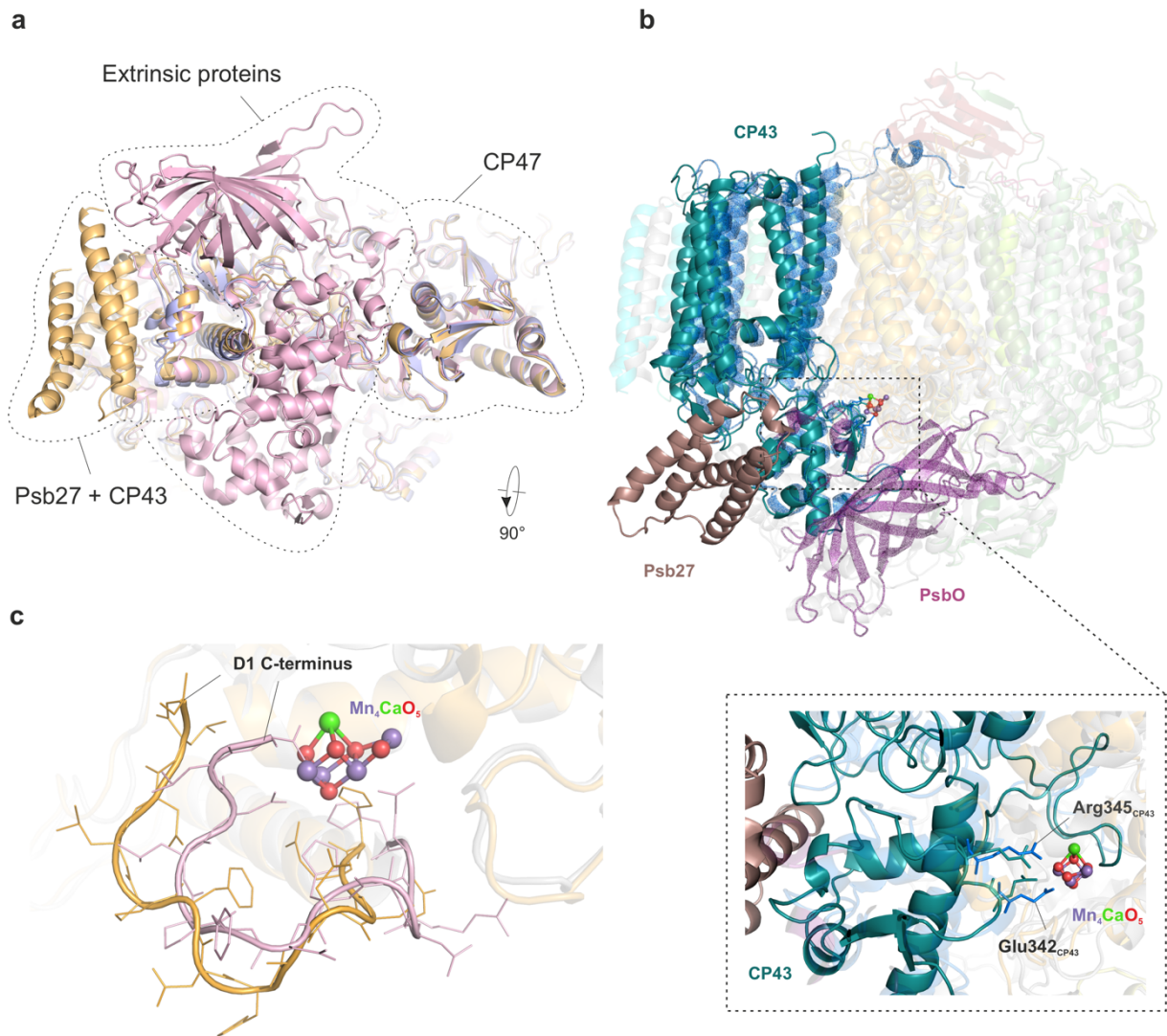
**Fig. 4. Structural changes of the D1 and D2 D-E loops induced by binding of Psb28 and Psb34.** **a**, Side view of the CP43 antenna protein in PSII-I (*teal*) and the PSII-M control (*light blue*). **b**, Structural changes between PSII-I and the PSII-M control in the cytoplasmic D2 D-E loop (*yellow*: 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  $Q_B$  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 Fig. S5c-h for enlarged views of the  $Q_A$  and  $Q_B$  binding site and the adjacent non-heme iron.



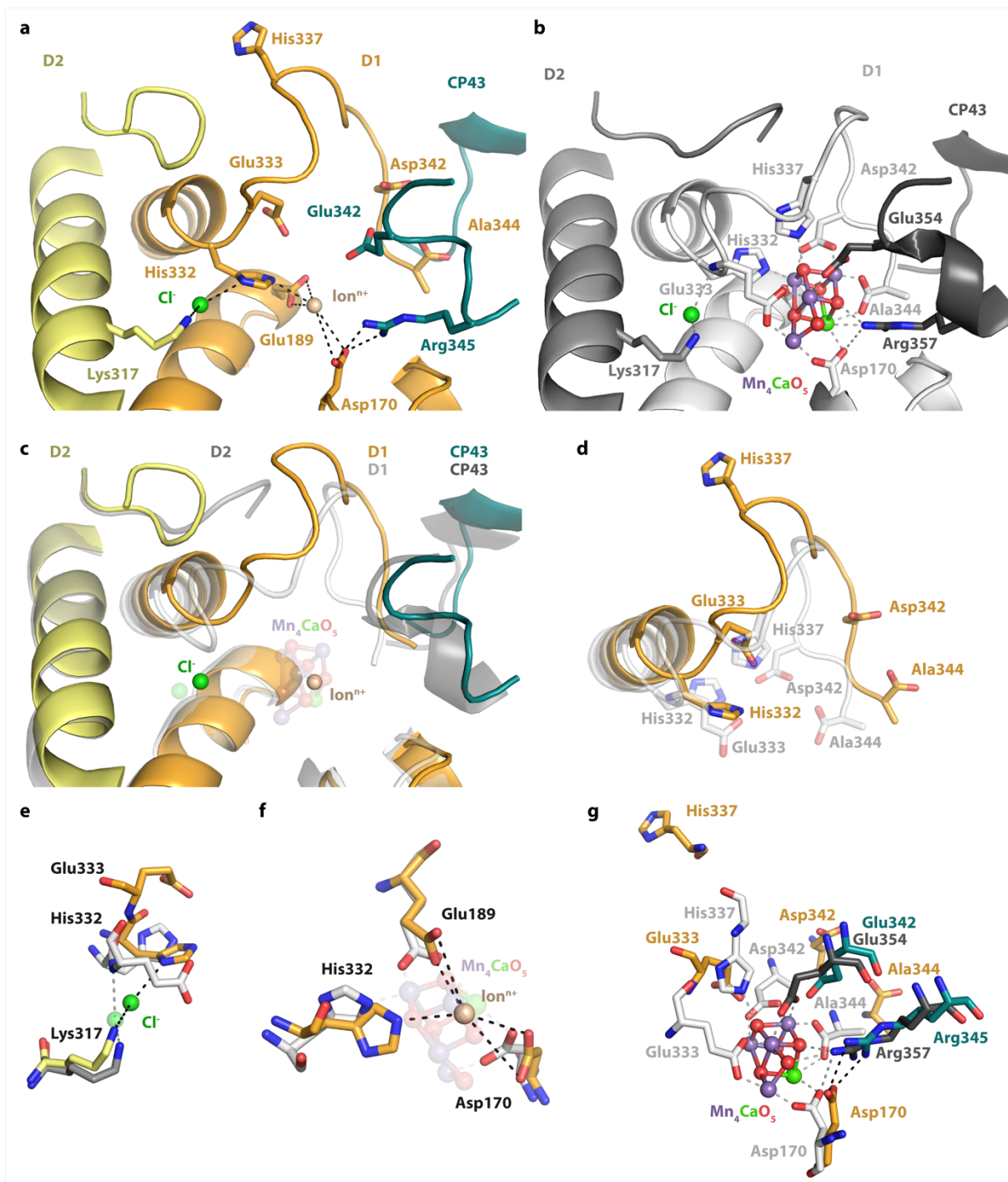
**Fig. 5: Binding of Psb28 displaces bicarbonate as a ligand of the non-heme iron and protects PSII from damage.** **a**, The electron transfer from  $PQ_A$  to  $PQ_B$  is coordinated by the non-heme iron ( $Fe^{2+}$ ), with the binding of bicarbonate (Bic) serving as a regulatory mechanism<sup>57</sup> 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<sup>69</sup> (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 evolving complex (OEC). **d**, Reoxidation of  $Q_A^-$  by direct and safe charge recombination is favored 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). Dotted corridors depict SD ( $n = 3$ ). **f**, The protective role of Psb28 binding was further confirmed by EPR spectroscopy using the spin probe TEMPD, which is specific for  $^1O_2$ , the major reactive oxygen species in PSII generated by triplet chlorophyll ( $^3P$ ).





**Fig. 6: The role of Psb27 in Mn<sub>4</sub>CaO<sub>5</sub> 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<sub>4</sub>CaO<sub>5</sub> cluster. We changed the numbering of CP43 residues due to a corrected N-terminal sequence ([www.UniProt.org](http://www.UniProt.org)). The residues correspond to Arg357 and Glu354 in previous publications. The high-resolution structure of the Mn<sub>4</sub>CaO<sub>5</sub> 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*).



**Figure 7: Conformational changes within the active site of the  $Mn_4CaO_5$  cluster.** The  $Mn_4CaO_5$  cluster performs PSII's unique water-splitting reaction. **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 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**),  $Ion^{n+}$  (**f**) and  $Mn_4CaO_5$  (**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.