

# Structures of the intermediates of Kok's photosynthetic water oxidation clock

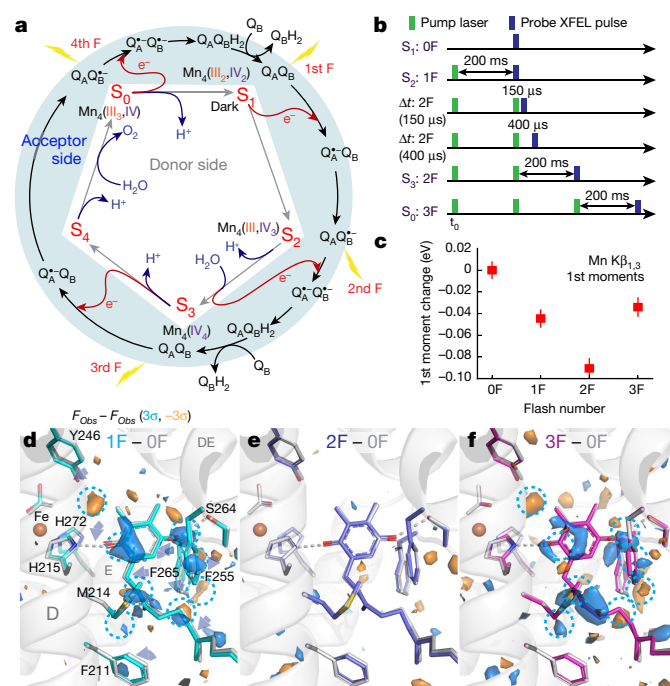
Jan Kern<sup>1</sup>, Ruchira Chatterjee<sup>1,12</sup>, Iris D. Young<sup>1,12</sup>, Franklin D. Fuller<sup>1,12</sup>, Louise Lassalle<sup>1</sup>, Mohamed Ibrahim<sup>2</sup>, Sheraz Gul<sup>1</sup>, Thomas Fransson<sup>3,11</sup>, Aaron S. Brewster<sup>1</sup>, Roberto Alonso-Mori<sup>4</sup>, Rana Hussein<sup>2</sup>, Miao Zhang<sup>2</sup>, Lacey Douthitt<sup>1</sup>, Casper de Lichtenberg<sup>5,6</sup>, Mun Hon Cheah<sup>6</sup>, Dmitry Shevela<sup>5</sup>, Julia Wersig<sup>2</sup>, Ina Seuffert<sup>2</sup>, Dimosthenis Sokaras<sup>7</sup>, Ernest Pastor<sup>1</sup>, Clemens Weninger<sup>4</sup>, Thomas Kroll<sup>7</sup>, Raymond G. Sierra<sup>4</sup>, Pierre Aller<sup>8</sup>, Agata Butryn<sup>8</sup>, Allen M. Orville<sup>8</sup>, Mengning Liang<sup>4</sup>, Alexander Batyuk<sup>4</sup>, Jason E. Koglin<sup>4</sup>, Sergio Carbajo<sup>4</sup>, Sébastien Boutet<sup>4</sup>, Nigel W. Moriarty<sup>1</sup>, James M. Holton<sup>1,7,9</sup>, Holger Dobbek<sup>2</sup>, Paul D. Adams<sup>1,10</sup>, Uwe Bergmann<sup>3</sup>, Nicholas K. Sauter<sup>1</sup>, Athina Zouni<sup>2\*</sup>, Johannes Messinger<sup>5,6\*</sup>, Junko Yano<sup>1\*</sup> & Vittal K. Yachandra<sup>1\*</sup>

Inspired by the period-four oscillation in flash-induced oxygen evolution of photosystem II discovered by Joliot in 1969, Kok performed additional experiments and proposed a five-state kinetic model for photosynthetic oxygen evolution, known as Kok's S-state clock or cycle<sup>1,2</sup>. The model comprises four (meta)stable intermediates ( $S_0$ ,  $S_1$ ,  $S_2$  and  $S_3$ ) and one transient  $S_4$  state, which precedes dioxygen formation occurring in a concerted reaction from two water-derived oxygens bound at an oxo-bridged tetra manganese calcium ( $Mn_4CaO_5$ ) cluster in the oxygen-evolving complex<sup>3–7</sup>. This reaction is coupled to the two-step reduction and protonation of the mobile plastoquinone  $Q_B$  at the acceptor side of PSII. Here, using serial femtosecond X-ray crystallography and simultaneous X-ray emission spectroscopy with multi-flash visible laser excitation at room temperature, we visualize all (meta)stable states of Kok's cycle as high-resolution structures (2.04–2.08 Å). In addition, we report structures of two transient states at 150 and 400 μs, revealing notable structural changes including the binding of one additional 'water', Ox, during the  $S_2 \rightarrow S_3$  state transition. Our results suggest that one water ligand to calcium (W3) is directly involved in substrate delivery. The binding of the additional oxygen Ox in the  $S_3$  state between Ca and Mn1 supports O–O bond formation mechanisms involving O5 as one substrate, where Ox is either the other substrate oxygen or is perfectly positioned to refill the O5 position during  $O_2$  release. Thus, our results exclude peroxo-bond formation in the  $S_3$  state, and the nucleophilic attack of W3 onto W2 is unlikely.

All four (meta)stable S-states of photosystem II (PSII) (Fig. 1a, Extended Data Fig. 1a) were populated by illumination of dark-adapted PSII crystals with 0, 1, 2 or 3 flashes (0F–3F; Fig. 1b). The approximately 2 Å resolution (Extended Data Fig. 1b–e, Extended Data Table 1) was sufficient for determining the positions of the oxygens bridging the metal atoms, in addition to the terminal water positions of the  $Mn_4CaO_5$  cluster; the former are critical for discriminating between proposed structures of the S-states, which was not possible in previous structures<sup>8,9</sup>. Reliable determination of the S-state composition of the PSII crystals obtained by each flash is essential for correct analysis of higher S-state structures. We therefore collected the Mn  $K\beta_{1,3}$  emission spectra in situ (see Methods)<sup>10</sup>, simultaneously with diffraction data. The first moment of the  $K\beta_{1,3}$  peak shifts towards lower energy in response to the first two flashes (0F→1F→2F) and to higher energy in the 3F sample (Fig. 1c), as expected from reported Mn redox states<sup>4,11–14</sup> (Fig. 1a). These data, in combination

with ex situ  $O_2$  evolution measurements, were used to determine the S-state distribution in each illuminated state (see Methods, Extended Data Fig. 2).

Isomorphous difference maps ( $F_{obs} - F_{obs}$ ) between the dark and flash-illuminated states at the acceptor side (Fig. 1d–f) indicate a clear



**Fig. 1 | The oxygen-evolving cycle in photosystem II.** **a**, Relationship between the redox chemistry at the donor (Kok's clock) and acceptor sides throughout the oxygen-evolving cycle. **b**, Relative timing of the visible laser (527 nm) pump and X-ray laser for the different illuminated states. **c**, First moment change of the Mn  $K\beta_{1,3}$  XES spectra from PSII crystals obtained in situ simultaneously with XRD. Data shown as mean  $\pm$  s.d. (see Methods). **d–f**,  $F_{obs} - F_{obs}$  isomorphous difference maps around plastoquinone  $Q_B$  contoured at  $+3\sigma$  (blue) and  $-3\sigma$  (orange) for the 1F, 2F and 3F states relative to 0F. The 0F stick model is shown in light grey; other models have carbons coloured cyan (1F), blue (2F) and magenta (3F).

<sup>1</sup>Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. <sup>2</sup>Institut für Biologie, Humboldt-Universität zu Berlin, Berlin, Germany.

<sup>3</sup>Stanford PULSE Institute, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. <sup>4</sup>LCLS, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. <sup>5</sup>Institutionen för Kemi, Kemiskt

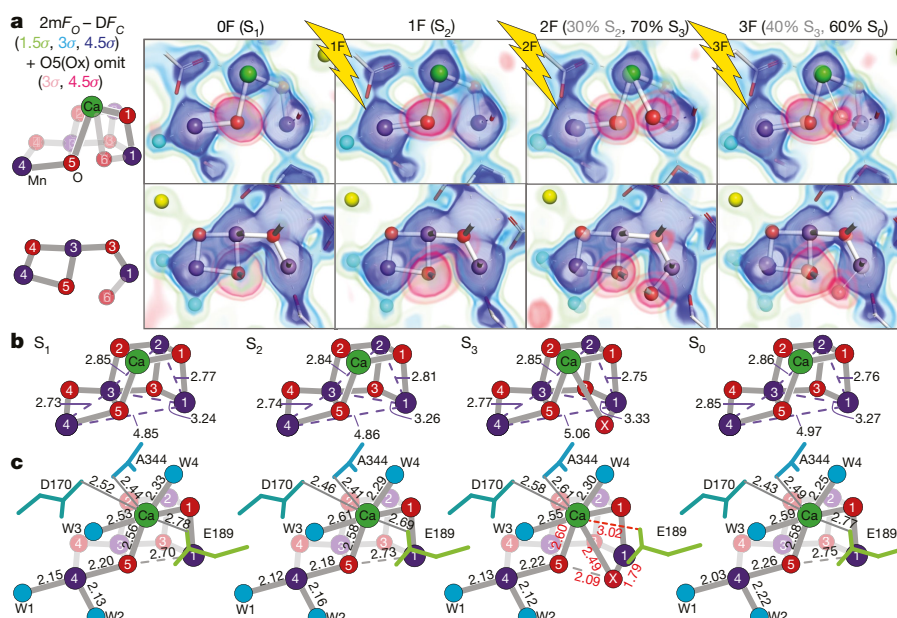
Biologiskt Centrum, Umeå Universitet, Umeå, Sweden. <sup>6</sup>Department of Chemistry—Ångström, Molecular Biomimetics, Uppsala University, Uppsala, Sweden. <sup>7</sup>SSRL, SLAC National Accelerator

Laboratory, Menlo Park, CA, USA. <sup>8</sup>Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, UK. <sup>9</sup>Department of Biochemistry and Biophysics, University of California, San

Francisco, CA, USA. <sup>10</sup>Department of Bioengineering, University of California Berkeley, Berkeley, CA, USA. <sup>11</sup>Present address: Interdisciplinary Center for Scientific Computing, University of

Heidelberg, Heidelberg, Germany. <sup>12</sup>These authors contributed equally: Ruchira Chatterjee, Iris D. Young, Franklin D. Fuller. \*e-mail: athina.zouni@hu-berlin.de; johannes.messinger@kemi.uu.se;

jjano@lbl.gov; vkyachandra@lbl.gov



**Fig. 2 | Stepwise changes at the OEC during the oxygen-evolving Kok cycle.** **a**, Left, labelled diagrams of the OEC atoms in two views. Right,  $2mF_{obs} - DF_{calc}$  density (green to blue) and O5/Ox omit map density (pink) shown as the overlay of several contour levels for the two views of the OEC in the 0F–3F states. The contributions of the S states to each data set for the two-component analysis are indicated in parentheses.

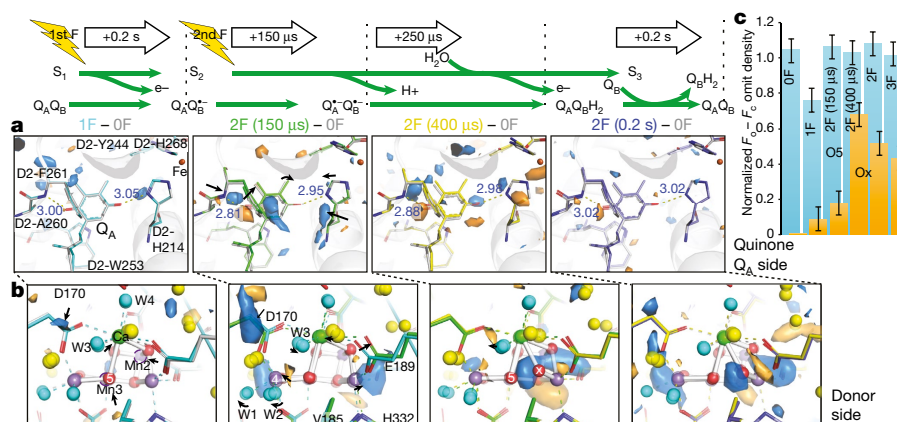
period-two oscillation of  $Q_B$  between its fully oxidized (0F, 2F) and semiquinone  $Q_B^{\bullet-}$  (1F) forms. A decrease in the  $B$ -factor of the  $Q_B$  site after 1F suggests that the quinone in the 1F sample is more tightly bound, owing to the formation of the semiquinone  $Q_B^{\bullet-}$ . The observation of  $Q_B^{\bullet-}$  in the 3F data implies that three electrons were successfully transferred from the  $Mn_4CaO_5$  cluster to the acceptor side, confirming S-state advancement in both PSII monomers (A and a; Fig. 1d–f, Extended Data Fig. 3).

In the  $S_1$  state, the cluster is in a distinct ‘right-open’ structure with no bond between Mn1 and O5; the Mn4–O5 distance is about 2.2 Å, whereas the Mn1–O5 distance is about 2.7 Å (Fig. 2; Extended Data Table 2), with Mn1 clearly pentacoordinate, and likely to be in the +3 oxidation state<sup>4,11,13–17</sup>. Open, non-cubane structures have been suggested by solution and single-crystal polarized extended X-ray

**b, c**, Atomic distances in the OEC in each S state in ångström, averaged across both monomers. Standard deviations for metal–metal, metal–bridging oxygen and metal–ligand distances are 0.1, 0.15 and 0.17 Å, respectively (see Methods). Ca remains 8-coordinate upon insertion of Ox by detaching D1-Glu189.

absorption fine structure (EXAFS) and electron paramagnetic resonance (EPR) data<sup>4,11,15,18</sup>.

Upon the transition from  $S_1$  to  $S_2$  (1F), one Mn is oxidized from +3 to +4. Figure 2a shows the electron density map ( $2mF_{obs} - DF_{calc}$ ) of the dark ( $S_1$ ) and 1F (predominantly  $S_2$ ) states. The structure of the cluster in the  $S_2$  state remains fundamentally unchanged, with the coordination numbers of all the metals preserved and in accordance with the similarity of the  $S_1$  and  $S_2$  EXAFS spectra<sup>4</sup>. This ‘right-open’ geometry is consistent with models of the  $S_2$  low spin ( $S_{total} = \frac{1}{2}$ ) configuration<sup>19</sup>, in which Mn4 is +4. There is no indication of a ‘left-open’ structure with an O5–Mn1 bond<sup>20</sup>. Only small structural changes are observed in the  $F_{obs} - F_{obs}$  difference map in the oxygen-evolving complex (OEC) region, with shifts of Ca, Mn3 and Mn2 and of residues Asp170 from subunit D1 and Glu354 from subunit CP43 of PS II



**Fig. 3 | The  $S_2 \rightarrow S_3$  transition in PSII.** **a**, **b**, Isomorphous difference density at quinone  $Q_A$  (**a**) and donor side (**b**) with the expected reduced state of  $Q_A$  at 2F (150 μs) and less pronounced at 2F (400 μs), and the oxidation of the OEC and insertion of Ox by 400 μs after the second flash.  $F_{obs} - F_{obs}$  difference densities between the various illuminated states and the 0F data are contoured at  $+3\sigma$  (blue) and  $-3\sigma$  (orange). The model for the 0F data is shown in light grey whereas carbons are coloured

in the models as follows: 1F (cyan), 2F (150 μs) (green), 2F (400 μs) (yellow) and 2F (0.2 s) (blue). **c**, Estimates of occupancies of O5 and Ox based on omit map peak heights normalized against the average electron density maximum at the O2 position in the omit maps. These match full occupancy of O5 throughout and Ox insertion by 400 μs after the second flash. Data shown as mean  $\pm$  s.d. based on the electron density value at the O2 position ( $n = 12$  observations, see Methods).

(Extended Data Fig. 4). Notably, the strongest difference is observed in the secondary coordination environment: a large negative peak at the location of water W20 (see below).

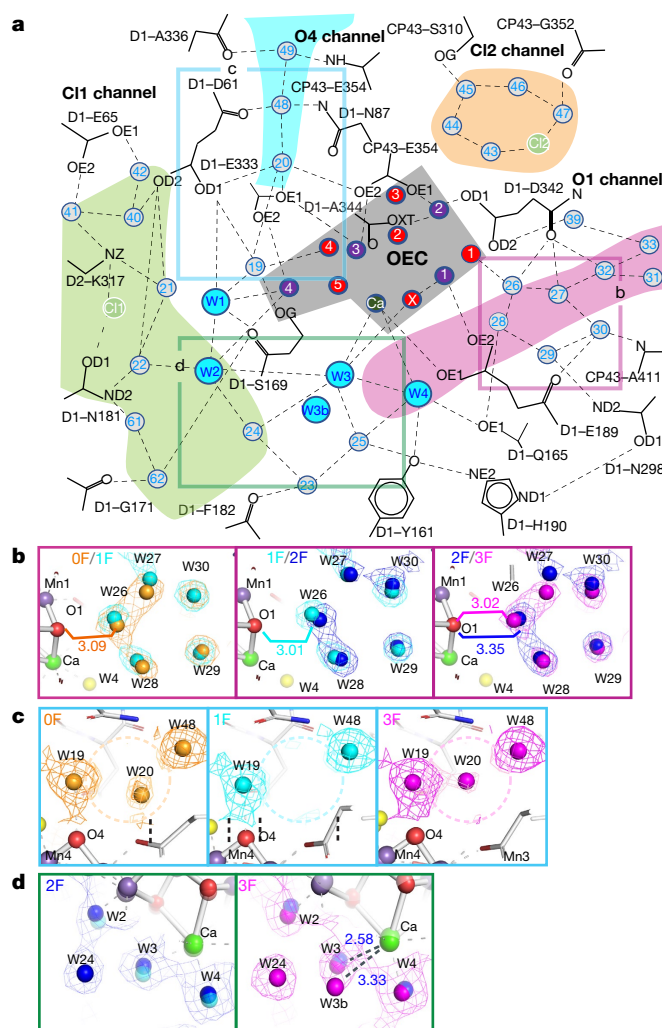
In Fig. 2a we evaluated the O5 position by overlaying the omit map on the  $2mF_{\text{obs}} - DF_{\text{calc}}$  map. In the 0F and 1F data, only one envelope of density is observed. Upon transition from  $S_2$  to  $S_3$  (2F), an additional feature appears near Mn1 that can be assigned to an inserted O atom (hydroxo or oxo). In the  $2mF_{\text{obs}} - DF_{\text{calc}}$  map of the 2F data, this additional density is visible as a small but distinct bulge that overlaps with the dominant Mn1 density (Extended Data Fig. 5). The presence of this additional density is distinguishable only in the omit map and in the  $2mF_{\text{obs}} - DF_{\text{calc}}$  map at high resolution, and not in the earlier 2.25 Å resolution map<sup>8</sup> (Extended Data Fig. 5).

We refined the 2F data against two partial-occupancy models at the catalytic site based on the S-state populations (see Methods). The S<sub>2</sub>-state structure, which is the refined structure for 1F, was fixed at 30% in the 2F state. The 2F data were then used to refine the S<sub>3</sub>-state structure at 70% occupancy. In the refined S<sub>3</sub> structure, the Mn1–Mn4 and Mn1–Mn3 distances are elongated by about 0.2 and 0.07 Å, respectively, relative to the S<sub>2</sub> structure, and the newly inserted hydroxide or oxo (Ox in Fig. 2a, b), located about 1.8 Å from Mn1, occupies its sixth coordination site. This change in coordination of Mn1 from 5- to 6-coordinate is in line with the proposed oxidation of Mn1 from +3 to +4 in the S<sub>2</sub>→S<sub>3</sub> transition<sup>21–23</sup>. Ox is also bound to Ca (2.50 Å) and it is closer to Ca than is O5 (2.60 Å). The Ca coordination number, however, remains eight, as D1–Glu189, which was ligated to Ca in the S<sub>1</sub> and S<sub>2</sub> states (at 2.78 and 2.69 Å, respectively), moves away from Ca in the S<sub>3</sub> state (3.01 Å), making space for Ox (Fig. 2c). These movements are accompanied by changes in the positions of nearby residues (His332, Glu333, His337, Asp342, Ala344, Asp170 of D1, and Glu354, Arg357 of CP43; Extended Data Fig. 4). We note that the current data cannot tell us whether Ox is oxo or hydroxo. However, the position of Ox is compatible with a deprotonated oxo bridge<sup>17</sup>.

Suga et al.<sup>9</sup> recently reported insertion of an O6 atom at 1.5 Å from O5 in their 2F sample (ex situ estimation 46% S<sub>3</sub> state) based on a 2F – 0F difference map at 2.35 Å resolution and proposed formation of a peroxide bond between O5 and O6 in the S<sub>3</sub> state. The O5–Ox distance of 2.1 Å in our data is about 0.6 Å longer than the O5–O6 bond modelled by Suga et al.<sup>9</sup>, and the location of Ox is 0.9–1.0 Å away from O6. Thus, our data do not agree with the formation of a peroxide-like bond in the S<sub>3</sub> state. We also note that if a peroxide-like bond were to form in the S<sub>3</sub> state, it would have to be accompanied by reduction of Mn, which conflicts with various spectroscopic observations<sup>4,15,21</sup> including our current in situ X-ray emission spectroscopy (XES) data (Fig. 1c), which show oxidation of Mn upon the S<sub>2</sub>→S<sub>3</sub> transition. We conclude that the differences in interpretation of the data between Suga et al.<sup>9</sup> and Young et al.<sup>8</sup> arose from uncertainty when determining oxygen positions at approximately 2.3 Å resolution, whereas our current data clearly show that Ox is bound to Mn1 and Ca, and that there is no peroxo-bond formation with O5 in the S<sub>3</sub> state.

The third laser flash (3F) advances the highest-oxidized metastable S<sub>3</sub> state to the most reduced S<sub>0</sub> state by releasing O<sub>2</sub> and acquiring one water molecule, resetting the catalytic cycle of Kok's clock. The S<sub>0</sub> structure shows the loss of Ox and the return to a motif similar to the dark stable S<sub>1</sub> state (Fig. 2b, c). The decrease in density for Ox in the 3F state (Fig. 2a) is in line with the 60% S<sub>0</sub> population (see Methods).

We collected data sets from two transient states at time points during the  $S_2 \rightarrow S_3$  transition, 150 and 400  $\mu\text{s}$  after the second flash (Fig. 1b), with resolutions of 2.5 and 2.2  $\text{\AA}$ , respectively, and computed  $F_{\text{obs}} - F_{\text{obs}}$  maps with the 0F data (Fig. 3, Extended Data Fig. 6, Supplementary Video 1). At the acceptor side, prominent difference peaks are visible in the vicinity of the primary quinone acceptor,  $Q_A$  (Fig. 3a), resembling the differences observed at the  $Q_B$  site upon formation of  $Q_B^{\bullet-}$  0.2 s after the first flash (Fig. 1d) and indicating formation of the reduced  $Q_A^{\bullet-}$  semiquinone. The same features are visible 250  $\mu\text{s}$  later in the 2F(400  $\mu\text{s}$ ) - 0F difference map, but at reduced intensity. By contrast, the corresponding  $F_{\text{obs}} - F_{\text{obs}}$  difference maps for the data sets collected



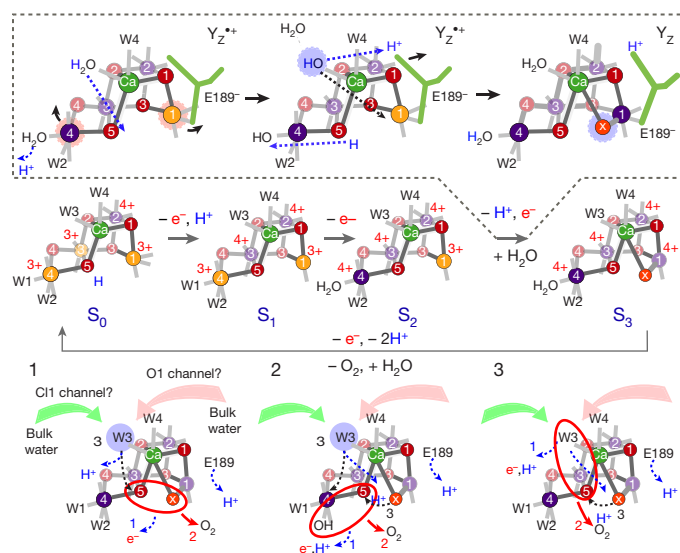
**Fig. 4 | Water network around the OEC.** **a**, Schematic of the H-bonding network surrounding the OEC indicating starting points of channels connecting the OEC to the solvent-exposed surface of PSII for possible water movement and proton transfer. **b–d**, Changes in selected water positions between the 0F and 3F states overlaid with  $2mF_{\text{obs}} - DF_{\text{calc}}$  maps in those states contoured at  $1.5\sigma$ . Positions in the schematic view are indicated by boxes of the same colour. Selected distances are given in ångström. **b**, Oscillation of the cluster of five waters next to O1 with the O1–O26 distance alternating long–short–long–short from 0F – 3F. **c**, W and its direct surroundings, indicating disappearance of W20 in the 1F ( $S_0$ ) state and reappearance in 3F ( $S_0$ ). **d**, Environment of Ca-bound W3, indicating the presence of a second water position for W3 in the  $S_0$  state (3F).

0.2 s after the first and second flashes do not show any indication of  $\text{Q}_\text{A}^{\bullet-}$  formation, following known kinetics for  $\text{Q}_\text{A}^{\bullet-}$  formation (sub-microsecond) and decay (on the order of hundreds of microseconds)<sup>24</sup>.

At the OEC, the  $2F(150\ \mu\text{s}) - 0F_{\text{obs}} - F_{\text{obs}}$  map shows that the first event after the absorption of the second photon is a movement of Mn4 and Mn1 away from each other by about  $0.2\ \text{\AA}$  (Fig. 3b). Density at the Ox site becomes visible in the  $2F(400\ \mu\text{s}) - 0F_{\text{obs}} - F_{\text{obs}}$  map as well as in the  $2mF_{\text{obs}} - DF_{\text{calc}}$  map, indicating the ligation of Ox to the Mn1 open coordination site. The Ox density (shown in Fig. 3c with the O5 density) increases substantially at  $400\ \mu\text{s}$ , and decreases in the 3F data. The remaining Ox density in the 3F data is explained by the approximately 40%  $S_2$  fraction.

It has been suggested that the  $S_2 \rightarrow S_3$  transition involves a proton transfer followed by an electron transfer<sup>7,25</sup>. We hypothesize that the proton transfer triggers the shifts of the Mn4 and Mn1 positions in the early stage of the  $S_2 \rightarrow S_3$  transition, including shifts of water positions (see below). One option is that W1 deprotonates, while W3 transfers





**Fig. 5 | Schematic structures of the S states in the Kok cycle of PSII and proposed reaction sequences for O–O bond formation in the  $S_4$  state.** The likely position of Mn oxidation states ( $Mn^{3+}$  is depicted in orange,  $Mn^{4+}$  in purple) as well as protonation and deprotonation reactions are indicated for each S state; the proposed steps in the  $S_2 \rightarrow S_3$  transition, including Ox insertion, are indicated in the dashed box with blue dashed arrows signifying atom movements. Three likely options (1, 2 and 3) for the final  $S_3 \rightarrow S_0$  transition are given in the bottom part, including possible order of 1) electron and proton release; 2) O–O bond formation and  $O_2$  release; and 3) refilling of the empty substrate site.

a proton to O5, weakening the O–Mn interaction and allowing the elongation of Mn1 and Mn4<sup>26</sup> that is necessary for subsequent oxygen insertion at Mn1, which is coupled to a proton transfer from O5 to W1. D1-Tyr161 ( $Y_Z$ ) is located about 4 Å from D1-Glu189. The formation of the positive charge at  $Y_Z$ /His189 that precedes oxidation of the OEC could trigger these structural changes in its surrounding, inducing a shift of D1-Glu189 away from Ca as observed in the 2F (150  $\mu$ s) data. The subsequent oxidation of Mn1 appears to be directly coupled to the insertion of Ox at the Mn1 open-coordination site observed in the 2F (400  $\mu$ s) data (Fig. 3b, c). We do not observe the formation of a ‘left-open’ structure that has been proposed on the basis of DFT-based studies for the early stage of the  $S_2 \rightarrow S_3$  transition<sup>20</sup>.

The OEC is embedded in an extended network of H bonds between amino acid residues and waters, which connects it to bulk water and is essential for its function (Fig. 4a, Extended Data Fig. 7, Extended Data Table 3). We observe several substantial changes in this network during the S-state cycle. Movements of the H-bonded water molecules W26–30 (Fig. 4b), of which W26 is located close to O1, indicate that these may be part of a  $H_2O/H^+$  transfer pathway or take part in charge redistributions within the  $Mn_4CaO_5$  cluster during the S-state cycle. W20 is lost in the  $S_1 \rightarrow S_2$  transition and reappears during the  $S_3 \rightarrow S_0$  transition (Fig. 4c, Extended Data Fig. 7, Supplementary Video 2), suggesting that the O4 channel may be used for proton release in the  $S_0 \rightarrow S_1$  transition but not in the  $S_2 \rightarrow S_3$  and  $S_3 \rightarrow S_0$  transitions. Suga et al.<sup>9</sup> did not report  $S_2$ -state data, but observed the loss of this water in the 2F data and related it to water insertion during  $S_3$ -state formation, which was further reinforced by the proposed computational ‘pivot’ or ‘carousel’ mechanisms involving the Mn4 site<sup>20,27</sup>. However, as this water is already absent in  $S_2$ , its loss may not have a direct relationship to the formation of  $S_3$ .

Additional water positions are observed near the Ca-bound W3 in the  $S_0$  structure: W3b at 3.25 Å to Ca has approximately 60% occupancy in monomer A (Fig. 4d, Extended Data Fig. 7), while the Ca-bound W3 is at 2.59 Å (approximately 40%). In monomer a, a smaller but similar positive peak is observed in the  $mF_{obs} - DF_{calc}$  map. Therefore, we hypothesize that W3 is the entrance site for the water, which is incorporated into the OEC during the  $S_3 \rightarrow S_0$  transition. As indicated

by the changes in the Ca ligation environment, W3 may play a similar role in the  $S_2 \rightarrow S_3$  transition. We therefore suggest that water at the W3 position may act as the entrance site or ‘parking place’ for either the substrate or the next substrate water in the  $S_2 \rightarrow S_3$  and  $S_3 \rightarrow S_0$  transitions, in agreement with earlier suggestions<sup>5,25,28–30</sup>. Possible access routes of water molecules to W3 are shown in Extended Data Fig. 7 (see also Fig. 4a).

Figure 5 summarizes the structures we determined for all of Kok’s S-states and our interpretation of the S-state-dependent changes with regard to the mechanism of water oxidation. As outlined above, we propose that the Ox site is filled by W3 during the  $S_2 \rightarrow S_3$  transition. The presence of Ox at Mn1, 2.1 Å from O5, suggests that Ox either forms the O–O bond with O5 in the  $S_3 \rightarrow S_0$  transition (case 1 in Fig. 5), or that Ox is placed between Ca and Mn1 to replace O5 during  $O_2$  formation or release. In the latter case, O–O bond formation may occur between O5 and W2 (2) or O5 and W3 (3). We exclude a nucleophilic attack mechanism of W3 or a protein-bound water molecule on W2 or W1, as the basis for these suggestions was that the  $Mn_3CaO_4$  cubane of the  $Mn_4CaO_5$  cluster acts as a ‘battery’ by storing oxidizing equivalents, but remains structurally unmodified in a closed cubane geometry<sup>31</sup>. This is in contrast to our data that show intricate structural changes in an open cubane involving ligand detachment and the formation of a new Ca–Ox–Mn1 bridge.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41586-018-0681-2>.

Received: 17 April 2018; Accepted: 22 August 2018;

Published online 7 November 2018.

- Kok, B., Forbush, B. & McGloin, M. Cooperation of charges in photosynthetic  $O_2$  evolution—I. A linear four step mechanism. *Photochem. Photobiol.* **11**, 457–475 (1970).
- Joliot, P., Barbieri, G. & Chabaud, R. A new model of photochemical centers in system-2. *Photochem. Photobiol.* **10**, 309–329 (1969).
- Hillier, W. & Wydrzynski, T.  $^{18}O$ -water exchange in photosystem II: substrate binding and intermediates of the water splitting cycle. *Coord. Chem. Rev.* **252**, 306–317 (2008).
- Yano, J. & Yachandra, V.  $Mn_4Ca$  cluster in photosynthesis: where and how water is oxidized to dioxygen. *Chem. Rev.* **114**, 4175–4205 (2014).
- Cox, N. & Messinger, J. Reflections on substrate water and dioxygen formation. *Biochim. Biophys. Acta* **1827**, 1020–1030 (2013).
- Debus, R. J. FTIR studies of metal ligands, networks of hydrogen bonds, and water molecules near the active site  $Mn_4CaO_5$  cluster in photosystem II. *Biochim. Biophys. Acta* **1847**, 19–34 (2015).
- Klauss, A., Haumann, M. & Dau, H. Seven steps of alternating electron and proton transfer in photosystem II water oxidation traced by time-resolved photothermal beam deflection at improved sensitivity. *J. Phys. Chem. B* **119**, 2677–2689 (2015).
- Young, I. D. et al. Structure of photosystem II and substrate binding at room temperature. *Nature* **540**, 453–457 (2016).
- Suga, M. et al. Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL. *Nature* **543**, 131–135 (2017).
- Kern, J. et al. Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science* **340**, 491–495 (2013).
- Kulik, L. V., Epel, B., Lubitz, W. & Messinger, J. Electronic structure of the  $Mn_4O_5Ca$  cluster in the  $S_0$  and  $S_2$  states of the oxygen-evolving complex of photosystem II based on pulse  $^{55}Mn$ -ENDOR and EPR spectroscopy. *J. Am. Chem. Soc.* **129**, 13421–13435 (2007).
- Dau, H. & Haumann, M. The manganese complex of photosystem II in its reaction cycle—basic framework and possible realization at the atomic level. *Coord. Chem. Rev.* **252**, 273–295 (2008).
- Peloquin, J. M. et al.  $^{55}Mn$  ENDOR of the  $S_2$ -state multiline EPR signal of photosystem II: Implications on the structure of the tetranuclear cluster. *J. Am. Chem. Soc.* **122**, 10926–10942 (2000).
- Krewald, V. et al. Metal oxidation states in biological water splitting. *Chem. Sci. (Camb.)* **6**, 1676–1695 (2015).
- Chernev, P. et al. Merging structural information from X-ray crystallography, quantum chemistry, and EXAFS spectra: The oxygen-evolving complex in PSII. *J. Phys. Chem. B* **120**, 10899–10922 (2016).
- Tanaka, A., Fukushima, Y. & Kamiya, N. Two different structures of the oxygen-evolving complex in the same polypeptide frameworks of photosystem II. *J. Am. Chem. Soc.* **139**, 1718–1721 (2017).
- Yamaguchi, K. et al. Theory of chemical bonds in metalloenzymes XXI. Possible mechanisms of water oxidation in oxygen evolving complex of photosystem II. *Mol. Phys.* **116**, 717–745 (2018).



18. Yano, J. et al. Where water is oxidized to dioxygen: structure of the photosynthetic Mn<sub>4</sub>Ca cluster. *Science* **314**, 821–825 (2006).
  19. Pantazis, D. A., Ames, W., Cox, N., Lubitz, W. & Neese, F. Two interconvertible structures that explain the spectroscopic properties of the oxygen-evolving complex of photosystem II in the S<sub>2</sub> state. *Angew. Chem. Int. Ed.* **51**, 9935–9940 (2012).
  20. Retegan, M. et al. A five-coordinate Mn(IV) intermediate in biological water oxidation: spectroscopic signature and a pivot mechanism for water binding. *Chem. Sci.* **7**, 72–84 (2016).
  21. Dau, H. & Haumann, M. Considerations on the mechanism of photosynthetic water oxidation — dual role of oxo-bridges between Mn ions in (i) redox-potential maintenance and (ii) proton abstraction from substrate water. *Photosynth. Res.* **84**, 325–331 (2005).
  22. Cox, N. et al. Photosynthesis. Electronic structure of the oxygen-evolving complex in photosystem II prior to O–O bond formation. *Science* **345**, 804–808 (2014).
  23. Siegbahn, P. E. M. Structures and energetics for O<sub>2</sub> formation in photosystem II. *Acc. Chem. Res.* **42**, 1871–1880 (2009).
  24. de Wijn, R. & van Gorkom, H. J. Kinetics of electron transfer from Q<sub>A</sub> to Q<sub>B</sub> in photosystem II. *Biochemistry* **40**, 11912–11922 (2001).
  25. Sakamoto, H., Shimizu, T., Nagao, R. & Noguchi, T. Monitoring the reaction process during the S<sub>2</sub>→S<sub>3</sub> transition in photosynthetic water oxidation using time-resolved infrared spectroscopy. *J. Am. Chem. Soc.* **139**, 2022–2029 (2017).
  26. Rossini, E. & Knapp, E.-W. Protonation equilibria of transition metal complexes: from model systems toward the Mn-complex in photosystem II. *Coord. Chem. Rev.* **345**, 16–30 (2017).
  27. Askerka, M., Wang, J., Vinyard, D. J., Brudvig, G. W. & Batista, V. S. S<sub>3</sub> state of the O<sub>2</sub>-evolving complex of photosystem II: insights from QM/MM, EXAFS, and femtosecond X-ray diffraction. *Biochemistry* **55**, 981–984 (2016).
  28. Bousac, A. & Rutherford, A. W. Nature of the inhibition of the oxygen-evolving enzyme of photosystem II induced by NaCl washing and reversed by the addition of Ca<sup>2+</sup> or Sr<sup>2+</sup>. *Biochemistry* **27**, 3476–3483 (1988).
  29. Tso, J., Sivaraja, M. & Dismukes, G. C. Calcium limits substrate accessibility or reactivity at the manganese cluster in photosynthetic water oxidation. *Biochemistry* **30**, 4734–4739 (1991).
  30. Suzuki, H., Sugiura, M. & Noguchi, T. Monitoring water reactions during the S-state cycle of the photosynthetic water-oxidizing center: detection of the DOD bending vibration by means of Fourier transform infrared spectroscopy. *Biochemistry* **47**, 11024–11030 (2008).
  31. Barber, J. A mechanism for water splitting and oxygen production in photosynthesis. *Nat. Plants* **3**, 17041 (2017).
- (J.Y., V.K.Y.), by National Institutes of Health (NIH) grants GM055302 (V.K.Y.), GM110501 (J.Y.) GM126289 (J.K.), GM117126 (N.K.S.), GM124149 and GM124169 (J.M.H.), the Ruth L. Kirschstein National Research Service Award (GM116423-02, F.D.F.), and Human Frontiers Science Project RGP0063/2013 (J.Y., U.B., A.Z.). We acknowledge the DFG-Cluster of Excellence “UniCat” coordinated by T.U. Berlin and Sfb1078, TP A5 (A.Z., H.D.); the Artificial Leaf Project (K&A Wallenberg Foundation 2011.0055) and Vetenskapsrådet (2016-05183) (J.M.); Diamond Light Source, Biotechnology and Biological Sciences Research Council (grant 102593) and a Strategic Award from the Wellcome Trust (A.M.O.). This research used NERSC, supported by DOE, under Contract No. DE-AC02-05CH11231. Synchrotron facilities at the ALS, Berkeley and SSRL, Stanford, were funded by DOE OBES. The SSRL Structural Molecular Biology Program is supported by the DOE OBER, and NIH (P41GM103393). LCLS and SSRL, SLAC National Accelerator Laboratory, are supported by DOE, OBES under Contract No. DE-AC02-76SF00515. We thank the staff at LCLS/SLAC and SSRL (BL 6-2, 7-3) and ALS (BL 5.01, 5.0.2, 8.2.1, 8.3.1).

**Author contributions** U.B., V.K.Y. and J.Y. conceived the experiment; R.A.-M., A.Z., J.M., U.B., N.K.S., J.K., V.K.Y. and J.Y. designed the experiment; R.C., M.I., L.L., R.H., M.Z., L.D., J.W., I.S., A.Z. and J.K. prepared samples; A. Batyuk, M.L., S.B., R.A.-M., J.E.K. and S.C. operated the MFX instrument; F.D.F., S.G., E.P., P.A., A.M.O., J.M. and J.K. developed, tested and ran the sample delivery system; M.H.C., D. Shevela, R.C., C.d.L., J.Y. and J.M. performed and analysed O<sub>2</sub> evolution and EPR measurements; R.C., F.D.F., S.G., M.I., C.d.L., M.H.C., I.D.Y., A.S.B., R.A.-M., R.H., M.Z., L.L., L.D., D. Sokaras, E.P., C.W., T.F., T.K., R.G.S., P.A., A. Butryn, A. Batyuk, M.L., S.B., J.E.K., S.C., A.M.O., A.Z., J.M., U.B., N.K.S., J.K., V.K.Y. and J.Y. performed the LCLS experiment; I.D.Y., A.S.B., N.W.M., J.M.H., P.D.A. and N.K.S. developed new software for data processing; I.D.Y., A.S.B., L.L., F.D.F., C.W., T.F., P.A., H.D., J.M.H., N.K.S. and J.K. processed and analysed XFEL data; J.M., J.K., V.K.Y. and J.Y. wrote the manuscript with input from all authors.

**Competing interests** The authors declare no competing interests.

#### Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41586-018-0681-2>.

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-018-0681-2>.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

**Correspondence and requests for materials** should be addressed to A.Z., J.M., J.Y. or V.K.Y.

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

**Acknowledgements** This work was supported by the Director, Office of Science, Office of Basic Energy Sciences (OBES), Division of Chemical Sciences, Geosciences, and Biosciences (CSGB), Department of Energy (DOE)

## METHODS

**Sample preparation.** PSII dimers were extracted and purified from *Thermosynechococcus elongatus* as reported previously<sup>32</sup>. PSII crystals ranging in size from 20 to 60  $\mu\text{m}$  were then prepared using a modified seeding protocol<sup>33</sup>. The crystals were dehydrated by treatment with high concentrations of PEG 5000. The final crystal suspension used for XRD measurements was in 0.1 M MES pH 6.5, 0.1 M ammonium chloride and 35% (w/v) PEG 5000, with  $\sim 0.5\text{--}0.8$  mM chlorophyll concentration. After loading into the sample delivery syringe (Hamilton gastight syringe, 1,000  $\mu\text{l}$ ), each sample syringe was given a preflash using green LED diodes (525 nm, Thorlabs, USA) to synchronize the samples, ensure that all centres have an oxidized tyrosine D and maximize yield of the higher S states upon subsequent flashing. The photon flux used at the sample for the preflash was  $2\ \mu\text{mol m}^{-2}\text{ s}^{-1}$ . The samples were exposed for ten seconds while rotating the syringe, and dark-adapted for 30 min to 2 h after the preflash. We note that the PSII core complexes in our sample preparation contain a sufficient number of natural quinones to drive the catalytic reaction through the cycle<sup>8,34,35</sup>.

**Characterization.** *O<sub>2</sub> activity.* Measurement of the O<sub>2</sub> yield by means of a Clark-type electrode under continuous illumination showed that the O<sub>2</sub> evolution rate of PSII solution before crystallization was  $2,300 \pm 100\ \mu\text{mol O}_2/(\text{mg(Chl)} \times \text{h})$  and after crystallization in the final buffer comprising 0.1 M MES pH 6.5, 0.1 M ammonium chloride it was  $2,100 \pm 80\ \mu\text{mol O}_2/(\text{mg(Chl)} \times \text{h})$  with 0.4 mM PPBQ. Approximately 90% activity was retained after crystallization.

*EPR measurements.* All sample batches used at the LCLS were checked with EPR for S-state turnover and Mn(II) content. The low-temperature X-band EPR spectra were measured using a Varian E109 EPR spectrometer equipped with a Model 102 Microwave bridge. For the turnover measurements, sample temperature was maintained at 8 K using an Air Products LTR liquid helium cryostat. The following spectrometer conditions were used: microwave frequency, 9.22 GHz; field modulation amplitude, 32 G at 100 kHz; microwave power, 20 mW. The turnover in the crystals was characterized using EPR spectroscopy. The turnover of the samples was measured as a function of the multiline EPR signal of the S<sub>2</sub> state, which oscillates with a period of four as a function of flash number. We observed the EPR signal for samples after a single turnover flash (1F) and monitored turnover by following decreases in amplitude of the multiline signal by subsequent flashes. In order to check the Mn(II) content, sample temperature was maintained at 20 K and the sensitivity of the measurement allowed to detect the presence of as low as 2% Mn(II) (compared to total Mn content) in the sample. The following spectrometer conditions were used: microwave frequency, 9.22 GHz; field modulation amplitude, 32 G at 100 kHz; microwave power, 1 mW. The Mn(II) content estimated by EPR agreed with that determined by in situ XES measurements<sup>36</sup>.

*Membrane inlet mass spectroscopy measurements.* The S-state advancement of crystals was also evaluated by membrane inlet mass spectroscopy (MIMS). A crystal suspension with approximately 10% <sup>18</sup>O-labelled water was placed in the thin-layer MIMS setup and subjected to a laser preflash before dark adaptation for 40 min at room temperature. The sample was then subjected to 2F at 5 Hz frequency and the O<sub>2</sub> yield was detected as a peak at  $m/z = 34$ . The procedure was repeated for 3F and 4F using new crystal suspensions each time. The O<sub>2</sub> yield pattern as function of flash number was calculated by subtracting the normalized O<sub>2</sub> yield of a flash number with the yield from the preceding flash number and normalized to 3F. The O<sub>2</sub> pattern can be fitted satisfactorily with an average miss parameter of 22%. The estimated S-state population is presented in Extended Data Fig. 2.

*Determination of the S-state population.* We evaluated the S-state advancement of PSII crystal suspension by two methods, in situ and ex situ. X-ray emission spectroscopy, which monitors the oxidation state of Mn, was collected simultaneously with the XRD data (Fig. 1c) as described previously<sup>10,36,37</sup>. The obtained single crystal XES spectra and details about XES data evaluation have been published recently<sup>36</sup>. Standard deviations for the first moments of the XES data shown as error bars in Fig. 1c were determined by random sampling of each of the data sets 1,000 times (each time randomly splitting the data into two subsets) and then calculating the standard deviation from the resulting 2,000 spectra for each flash state<sup>38</sup>.

Flash-induced oxygen measurements using MIMS for O<sub>2</sub> detection<sup>38–40</sup> were carried out before the XFEL experiment. We estimated the S-state population from both methods and these are presented in Extended Data Fig. 2. For the purpose of refinement of the structural models in the 2F data, we fixed the S<sub>3</sub> populations to 70% and S<sub>2</sub> to 30%. For the 3F data, we fixed the S<sub>3</sub> populations to 40%, and fit the remaining 60% as S<sub>0</sub>, because S<sub>0</sub> is the major population within the 60%. Further details regarding the refinement of the mixed models in the 2F and 3F data are given below (Model building and map calculation).

**Sample injection and illumination.** The crystallography data were collected at the MFX instrument of LCLS<sup>41,42</sup> during experiments LN84 and LQ39. The drop-on-tape (DOT) sample delivery method was used in combination with acoustic droplet ejection (ADE)<sup>37</sup>. For capturing the stable intermediates S<sub>2</sub>, S<sub>3</sub>, and S<sub>0</sub>, each droplet of the crystal suspension was illuminated by 120-ns laser pulses at 527 nm using a Nd:YLF laser (Evolution, Coherent) via fibre-coupled outputs 1, 2 and/or 3,

resulting in a delay time of 0.2 s between each illumination, and of 0.2 s between the last illumination and the X-ray probe<sup>37</sup>. We implemented a feedback control system of the belt speed and deposition delay, and the flashing delay and droplet phase were adjusted accordingly<sup>37</sup>.

For ex situ testing of light saturation for the DOT system, a 100–150  $\mu\text{m}$  thick sample film was established with the help of a washer between the silicon membrane of the mass spectrometer inlet and a thin microscope glass plate (thin layer MIMS setup). In this experiment, the samples were saturated at 70 mJ/cm<sup>2</sup>. The details have been described<sup>8</sup>. At the XFEL, a light intensity of  $120 \pm 10$  mJ/cm<sup>2</sup> was applied.

**X-ray diffraction setup and data processing.** PSII crystals were measured using X-ray pulses of  $\sim 40$  fs length at 9.5 keV and with an X-ray spot size at the sample of  $\sim 3\ \mu\text{m}$  in diameter. XRD data were collected on a Rayonix MX170 HS detector operating in the 2-by-2 binning mode at its maximum frame rate of 10 Hz. This mode provided the optimal trade-off of resolving power between adjacent Bragg reflections and quantity of images collected.

We developed the cctbx.xfel graphical user interface to track diffraction data acquisition, provide real-time feedback, and submit processing jobs. Processing jobs used `dials.stills_process`, a program within the cctbx.xfel framework that carries out lattice indexing, crystal model refinement, and integration and adopts a variety of defaults suited to XFEL still images<sup>43–48</sup>. For each image, strong spots are first selected. Next, candidate basis vectors describing the lattice of strong spots are identified, and an optional target cell is used to filter these candidates. A crystal model (composed of a unit cell and crystal orientation) is then refined to minimize differences between observed spot centroids and predicted positions, and this model is used to generate a complete set of indexed positions on the frame. Finally, signal at these positions is integrated and any corrections or uncertainties are taken into account. We found that with the stills-specific defaults and very few non-default parameters, 20–50% of shots (which we estimate to be the majority of the shots containing crystals) could be successfully indexed.

The powder diffraction pattern of a silver(I) behenate sample (Alfa Aesar) in a quartz capillary (Hampton Research, 10  $\mu\text{m}$  wall thickness) was used to obtain an initial estimate of detector distance. Initial indexing results were used to refine a detector distance and position for each interval between adjustments to the sample delivery system or detector position. These higher-precision detector positions were used in subsequent indexing and integration trials, resulting in a maximum of four distinct lattices indexed on a single shot.

`Cluster.unit_cell`, a command line tool in cctbx that clusters similar unit cells according to the Andrews–Bernstein distance metric<sup>49,50</sup>, was used to obtain the average unit cell. This unit cell was used as the target unit cell when reprocessing all experimental data with `dials.stills_process`.

A total of 1,565,863 integrated lattices were obtained using `dials.stills_process` with a target unit cell of  $a = 117.5\ \text{\AA}$ ,  $b = 222.8\ \text{\AA}$ ,  $c = 309.6\ \text{\AA}$ ,  $\alpha = \beta = \gamma = 90^\circ$  and the space group  $P2_12_12_1$ . Signal was integrated to the edges of the detector in anticipation of a per-image resolution cutoff during the merging step. Integrated intensities were corrected for absorption by the Kapton conveyor belt to match the position of the belt and crystals relative to the X-ray beam<sup>37</sup>.

Finally, XES data collected simultaneously with the diffraction images were used to sort out and exclude any sample batches that indicated the presence of Mn(II) released during the on-site crystallization<sup>36</sup>. It was also used to confirm the advancement of S states by fibre-coupled lasers and a free space laser.

Image sets were also culled to include only images diffracting beyond 6  $\text{\AA}$  (for small data sets) or beyond 3  $\text{\AA}$  (all others), similar to a procedure that has been used previously<sup>51</sup> to improve statistics in large data sets suffering from contamination by low-quality images. Until the experiments described here, we were typically data-limited and have focused data processing methods development on discovering how to extract the most signal from low-multiplicity data sets<sup>52,53</sup>. Several data visualization tools we implemented in the cctbx.xfel graphical user interface have made it possible to tune crystallization conditions and the sample delivery system to optimize diffraction quality early in an XFEL diffraction experiment, resulting in collection of much larger quantities of data. Although post-refinement and the per-image resolution cut-offs used in `cx.merge` downweigh or remove most spot predictions without signal, we still observed improvement in merging statistics and map quality when excluding lattices diffracting to a resolution poorer than 3  $\text{\AA}$  from these larger data sets, probably because of the limitation in orientational precision when indexing the small number of reflections visible on low-resolution stills.

The remaining integrated images were merged using `cx.merge` as described previously<sup>8</sup>, with a couple of modifications. The default unit cell outlier rejection mechanism in `cx.merge` was sufficiently selective on the image set curated as described above, so a pre-filtering step was not necessary. Also, a reference model and data set with a compatible unit cell—used by `cx.merge` during scaling—were available from previous beam times, so a preliminary merging step with PRIME was not necessary.

Final merged data sets were acquired for the 0F, 1F, 2F (150  $\mu$ s), 2F (400  $\mu$ s), 2F, and 3F states to resolutions between 2.50 and 2.04 Å, containing between 4,231 and 30,366 images (Extended Data Table 1). Additionally, data from all illuminated states were aggregated, culled to the subset of images extending past 2.2 Å, and merged as a separate 'combined' data set to 1.98 Å (results not shown).

**Model building and map calculation.** Initial structure refinement against the 'combined' data set at 1.98 Å was carried out starting from a previously acquired high-resolution PSII structure in the same unit cell (PDB ID: 5TIS) using phenix.refine<sup>54,55</sup>. After an initial rigid body refinement step, xyz coordinates and isotropic B-factors were refined for tens of cycles with automatic water placement enabled. Custom bonding restraints were used for the OEC (with large sigma values, to reduce the effect of the strain at the OEC on the coordinate refinement), chlorophyll-*a* (CLA, to allow correct placement of the Mg relative to the plane of the porphyrin ring), and unknown lipid-like ligands (STE). Custom coordination restraints overrode van der Waals repulsion for coordinated chlorophyll Mg atoms, the non-haem iron, and the OEC. Following real space refinement in Coot<sup>56</sup> of selected individual sidechains and the PsbO loop region and placement of additional water molecules, the model was refined for several additional cycles with occupancy refinement enabled, then as before without automatic water placement, and then as before with hydrogen atoms. NHQ flips and automatic linking were disabled throughout. A final 'combined' data set model was obtained with  $R_{\text{work}}/R_{\text{free}}$  of 17.92%/22.01%.

The above model was subsequently refined against the illuminated data sets to produce models that differed primarily at the OEC and plastoquinone, as confirmed by isomorphous difference maps, with the lattermost refinement settings and different OEC bonding restraints. OEC bonding restraints for the 0F data set prevented large deviations from the high-resolution dark state OEC structure reported by Suga et al. (PDB ID: 4UB6)<sup>53</sup>. Bonding restraints for the other data sets loosely restrained the models to metal–metal distances matching spectroscopic data and metal–oxygen distances matching the most likely proposed models<sup>57–61</sup>. A number of ordered water positions were excluded from subsequent automatic water placement rounds by renaming the residue names to OOO and supplying Phenix with a bonding restraint CIF dictionary for OOO identical to that for HOH, and the waters coordinating the OEC were incorporated into the OEC restraint CIF file directly. After 12–15 cycles of refinement in this manner, individual illuminated states at various resolutions were obtained ranging in  $R_{\text{work}}/R_{\text{free}}$  from 16.69%/24.60% to 19.33%/26.39% (Extended Data Table 1).

After the first cycles of refinement for the 2F data using the initial OEC model from the 0F data as starting point, a positive peak in the  $mF_{\text{obs}} - DF_{\text{calc}}$  density close to Mn1 became visible (see also Extended Data Fig. 5) and the automatic water placement step in refinement placed a water at the OEC between O5 and Mn1. We designated this new water as a potential Ox and added it to the OEC. We tested several starting positions for this Ox, including a position similar to O6 (Suga et al.)<sup>9</sup> at 1.5 Å from O5, but refinement resulted in shifts of the Ox away from O5 and close to Mn1.  $mF_{\text{obs}} - DF_{\text{calc}}$  difference maps calculated for different positions of this additional oxygen also confirmed a placement at about 1.8 Å from Mn1 and 2.1 Å from O5. For the final refinement, Ox was included in the CIF restraints for the OEC in the  $S_3$  state.

To best approximate the contributions of dimers that did not advance to the next S state owing to illumination misses, for the 2F and 3F data sets, the 2F and 3F models were split into A and B alternate conformers in regions of chains A/a, C/c and D/d surrounding (and including) the OEC. These residues are A55–65, A160–190, A328–344, C328, C354–358 and D352. Population of the  $S_3$  and  $S_0$  states in the 2F and 3F data was estimated on the basis of oxygen evolution and XES measurements (Extended Data Fig. 2) and rounded to the nearest 10%, yielding 70%  $S_3$  state population in the 2F data set and 60%  $S_0$  state population in the 3F data set. Accordingly, for the 2F data set, the main conformer across this entire region was set at 0.7 occupancy, and the minor conformer was set at 0.3 occupancy. Analogously, for the 3F data set, conformers were set at 0.6 and 0.4 to match an estimated 40% contribution from the  $S_3$  state and modelling the remaining 60% as the  $S_0$  state. The major conformer was allowed to refine as usual, while the minor conformer was fixed during refinement and set to match the major conformer of the previous S state (for example, fixed coordinates for  $S_2$  at 30% for the 2F,  $S_3$ -enriched state). This was achieved by least-squares fitting the refined model of the previous S state onto the new model at the split region in PyMol<sup>62</sup> and replacing the minor conformer atomic coordinates with the fitted model coordinates, then excluding the newly placed atoms from refinement in phenix.

Although phenix.refine supports modelling of three or more conformers, we limited our analysis to two conformers in consideration of both the limits of the resolution and the precision of the S-state contribution estimates, and we did not model a ~10% contribution of the  $S_1$  state in the 1F data set. When placing the refined  $S_3$  state model into the 3F model at 0.4 occupancy, we used the 0.7 occupancy model refined as described above, not the combination of both conformers. This analysis was not possible for the 2F (150  $\mu$ s) and 2F (400  $\mu$ s) models because

the complementary components would be time points 150  $\mu$ s and 400  $\mu$ s after the  $S_1$  state, structures we have not probed.

**Estimated positional precision.** The maximum-likelihood coordinate error calculated during refinement is a general-purpose metric for positional error but is subject to several limitations, including the impact of bonding, angle, coordination and other restraints on the refined model. Previously, we have generated ballpark estimates of positional error for various sections of the model by setting them to zero occupancy, conducting simulated annealing followed by refinement, placing the same components into omit density, and reporting the final magnitude of the shift between the centres of mass of the original and omit density-fitted components<sup>8</sup>. This is impractical for more than a handful of representative cofactors or segments of the main chain and is difficult for much smaller groups of atoms or individual atoms whose environments easily fill the missing density region if it is not artificially held open. We also tried multi-start kicked model refinement and found that the OEC refined back to nearly the same position in all trials. We therefore shifted our focus to a tool that perturbs the structure factors directly. By perturbing structure factors by  $\pm|F_{\text{obs}} - F_{\text{model}}|$  in 100 trials using the END/RAPID command line tools, we added noise proportional to the error in the model to generate 100 perturbed data sets for each illumination state, re-refined kicked models against each new data set, and calculated the mean and s.d. of selected bond distances across the re-refined models<sup>63</sup>. Metal–metal distances at the OEC had standard deviations between 0.08 and 0.12 Å across these trials, while distances between OEC metals and bridging oxygen atoms had standard deviations varying between 0.10 and 0.23 Å and distances between OEC metals and coordinating ligands were found to have standard deviations between 0.13 and 0.20 Å.

**Estimating the uncertainty of the omit densities.** Changes in the electron density at the positions of Ox and O5 were obtained from O5 and Ox omit maps and normalized against the average electron density maximum at the O2 position in O2 omit maps, assuming that O2 is always fully occupied in the different flash states. The standard deviation of the electron density value at O2 over all data sets and both monomers was also used to estimate the uncertainty of the normalized omit density. The omit densities of a particular data set were divided by the average omit densities of the chloride ions of the same data set to equate the densities from different data sets.

**Isomorphous difference maps.** Slight, nonphysical differences in merged unit cells were modelled across the illumination states in this sequence of data sets. Large distributions of unit cells derived from indexing with dials.stills\_process are known to reflect uncertainty in the crystal model, not variation among actual crystals<sup>64</sup>, and the distributions also shifted as sample-to-detector distance changed over the course of an LCLS shift. Because it was not possible to cycle through all illumination conditions throughout each of the experiments, the average unit cell dimensions varied across data sets as well, resulting in the aforementioned non-physical differences in merged unit cells. To obtain isomorphous difference maps without artefacts from these apparent unit cell variations we computed a second set of data, selecting only lattices with unit cells within 1% of the target unit cell. The resulting smaller data sets were at slightly reduced resolution (Extended Data Table 1b). However, because the merged unit cell differences were small, artefact free isomorphous difference maps could be calculated between pairs of these data sets. The corresponding .mtz and .pdb files for these smaller data sets are available from the authors upon request.

**Code availability.** The open source programs dials.stills\_process, the cctbx.xfel GUI and cxi.merge are distributed with DIALS packages available at <http://dials.github.io>, with further documentation available at <http://cci.lbl.gov/xfel>.

**Reporting summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

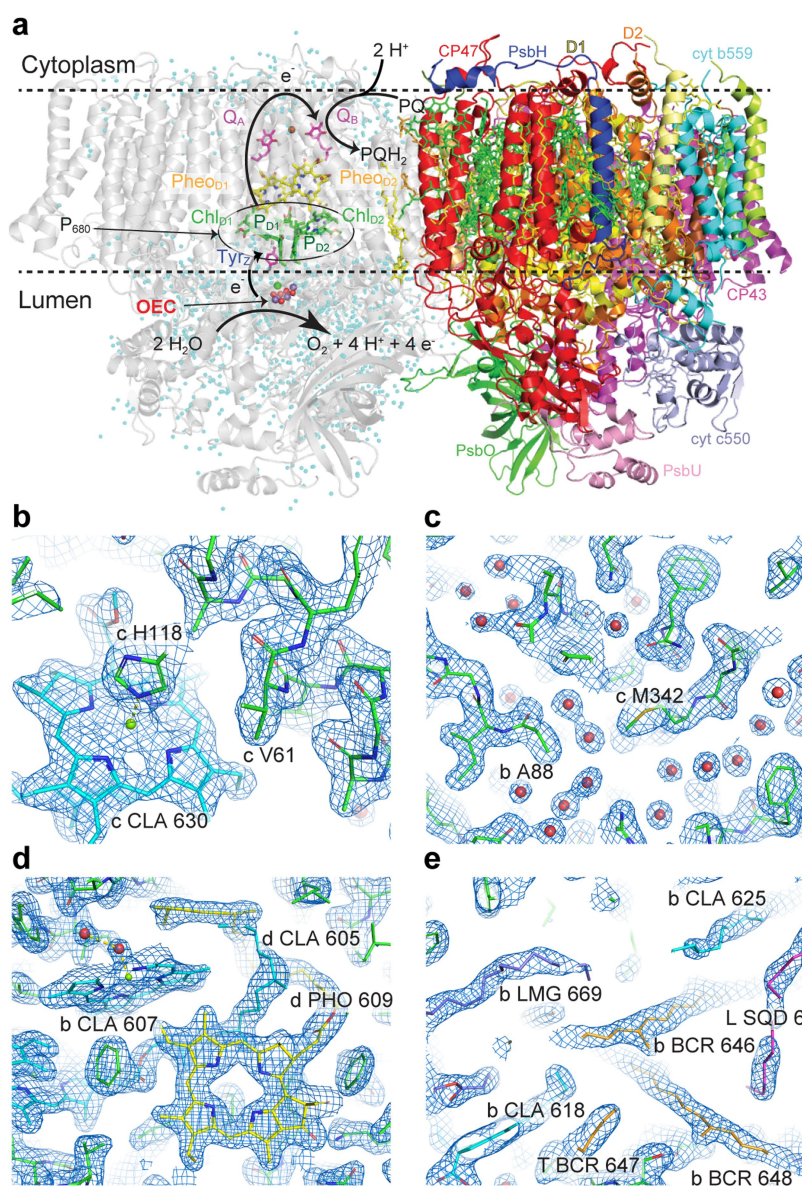
## Data availability

The atomic coordinates and structure factors have been deposited in the Protein Data Bank under the following accession codes: 6DHE for the 0F, 6DHF for the 1F, 6DHO for the 2F, 6DHG for the 2F(150 $\mu$ s), 6DHH for the 2F(400 $\mu$ s) and 6DHP for the 3F data.

- Hellmich, J. et al. Native-like photosystem II superstructure at 2.44 Å resolution through detergent extraction from the protein crystal. *Structure* **22**, 1607–1615 (2014).
- Ibrahim, M. et al. Improvements in serial femtosecond crystallography of photosystem II by optimizing crystal uniformity using microseeding procedures. *Struct. Dyn.* **2**, 041705 (2015).
- Guskov, A. et al. Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. *Nat. Struct. Mol. Biol.* **16**, 334–342 (2009).
- Krivanek, R., Kern, J., Zouni, A., Dau, H. & Haumann, M. Spare quinones in the Q<sub>B</sub> cavity of crystallized photosystem II from *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta* **1767**, 520–527 (2007).

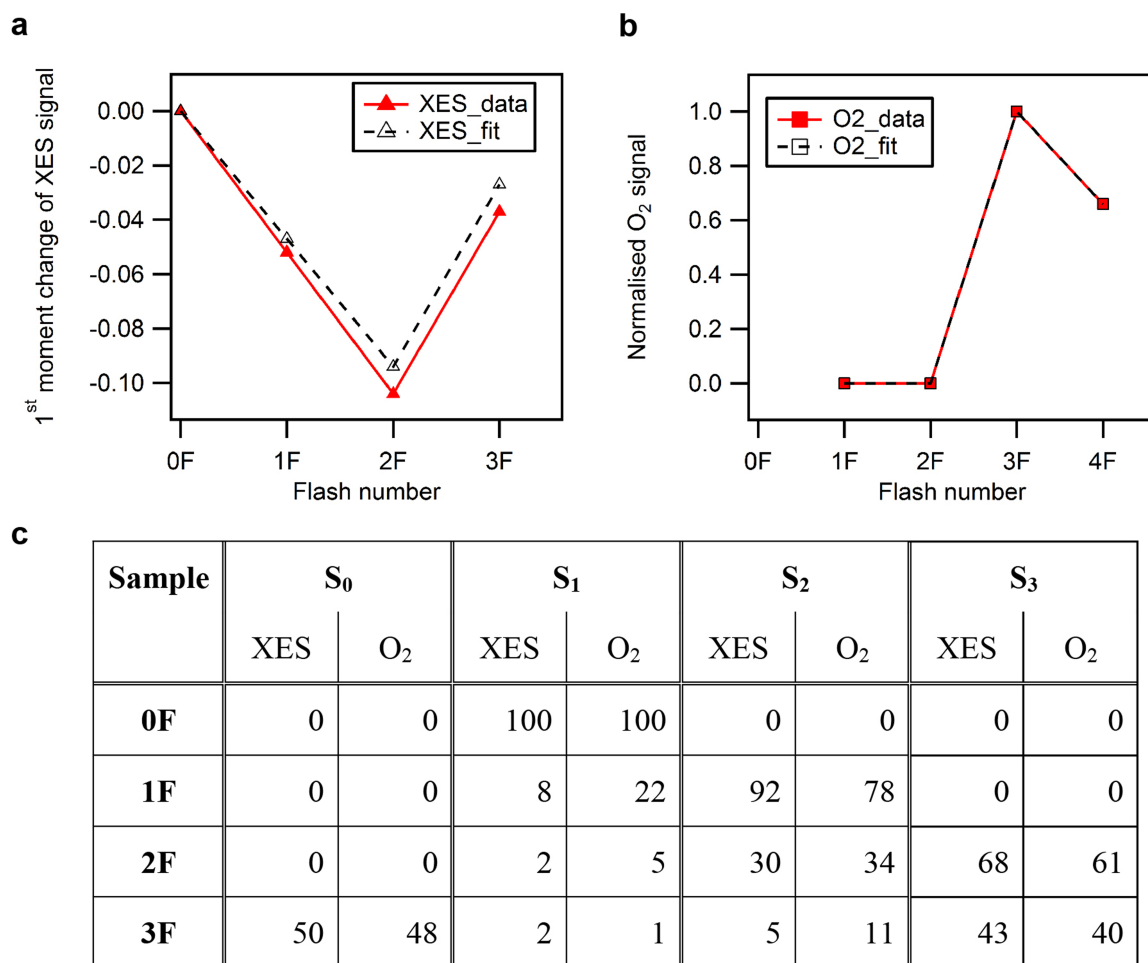


36. Fransson, T. et al. X-ray emission spectroscopy as an *in situ* diagnostic tool for X-ray crystallography of metalloproteins using an X-ray free-electron laser. *Biochemistry* **57**, 4629–4637 (2018).
37. Fuller, F. D. et al. Drop-on-demand sample delivery for studying biocatalysts in action at X-ray free-electron lasers. *Nat. Methods* **14**, 443–449 (2017).
38. Kern, J. et al. Taking snapshots of photosynthetic water oxidation using femtosecond X-ray diffraction and spectroscopy. *Nat. Commun.* **5**, 4371 (2014).
39. Yano, J. et al. in *Sustaining Life on Planet Earth: Metalloenzymes Mastering Dioxide and Other Chewy Gases, Metal Ions in Life Sciences* (eds Kroneck, P. M. H. & Sosa Torres, M. E.) 13–43 (Springer International Publishing, 2015).
40. Beckmann, K., Messinger, J., Badger, M. R., Wydrzynski, T. & Hillier, W. On-line mass spectrometry: membrane inlet sampling. *Photosynth. Res.* **102**, 511–522 (2009).
41. Emma, P. et al. First lasing and operation of an Ångström-wavelength free-electron laser. *Nat. Photon.* **4**, 641–647 (2010).
42. Boutet, S., Cohen, A. & Wakatsuki, S. The new macromolecular femtosecond crystallography (MX) instrument at LCLS. *Synchr. Radiat. News* **29**, 23–28 (2016).
43. Sauter, N. K. XFEL diffraction: developing processing methods to optimize data quality. *J. Synchr. Radiat.* **22**, 239–248 (2015).
44. Winter, G. et al. DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. D Struct. Biol.* **74**, 85–97 (2018).
45. Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. New Python-based methods for data processing. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1274–1282 (2013).
46. Hattne, J. et al. Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nat. Methods* **11**, 545–548 (2014).
47. Sauter, N. K. et al. Improved crystal orientation and physical properties from single-shot XFEL stills. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 3299–3309 (2014).
48. Waterman, D. G. et al. Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr. D Struct. Biol.* **72**, 558–575 (2016).
49. Zeldin, O. B. et al. Data Exploration Toolkit for serial diffraction experiments. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 352–356 (2015).
50. Andrews, L. C. & Bernstein, H. J. The geometry of Niggli reduction: BGAOL - embedding Niggli reduction and analysis of boundaries. *J. Appl. Crystallogr.* **47**, 346–359 (2014).
51. Suga, M. et al. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99–103 (2015).
52. Uervirojnangkoorn, M. et al. Enabling X-ray free electron laser crystallography for challenging biological systems from a limited number of crystals. *eLife* **4**, e05421 (2015).
53. Lyubimov, A. Y. et al. Advances in X-ray free electron laser (XFEL) diffraction data processing applied to the crystal structure of the synaptotagmin-1/SNARE complex. *eLife* **5**, e18740 (2016).
54. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
55. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
56. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
57. Wieghardt, K. The active-sites in manganese-containing metalloproteins and inorganic model complexes. *Angew. Chem. Int. Edn Engl.* **28**, 1153–1172 (1989).
58. Cinco, R. M. et al. Comparison of the manganese cluster in oxygen-evolving photosystem II with distorted cubane manganese compounds through X-ray absorption spectroscopy. *Inorg. Chem.* **38**, 5988–5998 (1999).
59. Mukhopadhyay, S., Mandal, S. K., Bhaduri, S. & Armstrong, W. H. Manganese clusters with relevance to photosystem II. *Chem. Rev.* **104**, 3981–4026 (2004).
60. Law, N. A., Caudle, M. T. & Pecoraro, V. L. Manganese redox enzymes and model systems: Properties, structures, and reactivity. *Adv. Inorg. Chem.* **46**, 305–440 (1999).
61. Tsui, E. Y., Kanady, J. S. & Agapie, T. Synthetic cluster models of biological and heterogeneous manganese catalysts for O<sub>2</sub> evolution. *Inorg. Chem.* **52**, 13833–13848 (2013).
62. Schrödinger LLC. The pymol molecular graphics system, version 1.8. (2015).
63. Lang, P. T., Holton, J. M., Fraser, J. S. & Alber, T. Protein structural ensembles are revealed by redefining X-ray electron density noise. *Proc. Natl Acad. Sci. USA* **111**, 237–242 (2014).
64. Brewster, A. S. et al. Improving signal strength in serial crystallography with DIALS geometry refinement. *Acta Crystallogr. D Struct. Biol.* **74**, 877–894 (2018).
65. Kupitz, C. et al. Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature* **513**, 261–265 (2014).
66. Ho, F. M. & Styring, S. Access channels and methanol binding site to the CaMn<sub>4</sub> cluster in Photosystem II based on solvent accessibility simulations, with implications for substrate water access. *Biochim. Biophys. Acta* **1777**, 140–153 (2008).
67. Vassiliev, S., Zaraiskaya, T. & Bruce, D. Exploring the energetics of water permeation in photosystem II by multiple steered molecular dynamics simulations. *Biochim. Biophys. Acta* **1817**, 1671–1678 (2012).
68. Murray, J. W. & Barber, J. Structural characteristics of channels and pathways in photosystem II including the identification of an oxygen channel. *J. Struct. Biol.* **159**, 228–237 (2007).
69. Gabdulkhakov, A. et al. Probing the accessibility of the Mn<sub>4</sub>Ca cluster in photosystem II: channels calculation, noble gas derivatization, and cocrystallization with DMSO. *Structure* **17**, 1223–1234 (2009).
70. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55–60 (2011).
71. Sakashita, N., Watanabe, H. C., Ikeda, T. & Ishikita, H. Structurally conserved channels in cyanobacterial and plant photosystem II. *Photosynth. Res.* **133**, 75–85 (2017).



**Extended Data Fig. 1 | Overview of the PSII structure and electron density maps of the 3F state.** **a**, Structure of the native PSII homodimer. In the left monomer the location of cofactors for the initial charge separation (P<sub>680</sub>, Pheo<sub>D1</sub>), and for the electron transfer leading to the reduction of the plastoquinone (Q<sub>A</sub>, Q<sub>B</sub>) at the acceptor side and to the oxidation of the OEC at the donor side by P680<sup>+</sup> are indicated. In the right monomer, the

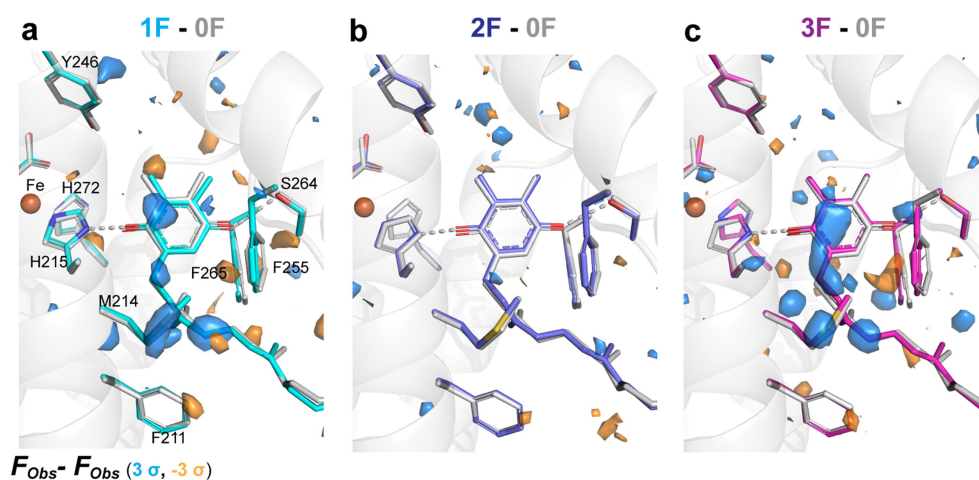
locations of the protein subunits are displayed. **b–d**,  $2mF_{\text{obs}} - DF_{\text{calc}}$  map (blue,  $1.5\sigma$  contour) obtained from the room temperature 3F data set. **b**, Density around the main chain and a chlorophyll. **c**, Well-resolved ordered water molecules. **d**, Chlorophyll and pheophytin molecules with well-resolved tails. **e**, Clear density in hydrophobic regions and along cofactor hydrocarbon tails.



**Extended Data Fig. 2 | Flash-induced S-state turnover of PSII micro crystals.** **a**, Change of the first moment of the in situ-measured Mn K $\beta$  XES as a function of flashes and fit to the data. **b**, Flash-induced O<sub>2</sub> yield as measured by MIMS as a function of flash number and fit to the data. **c**, The estimated S-state population (%) for each of the flash states from fitting of the XES data and of the flash-induced O<sub>2</sub> evolution pattern of a suspension of PSII crystals at pH 6.5. Two different fits were performed: a global fit

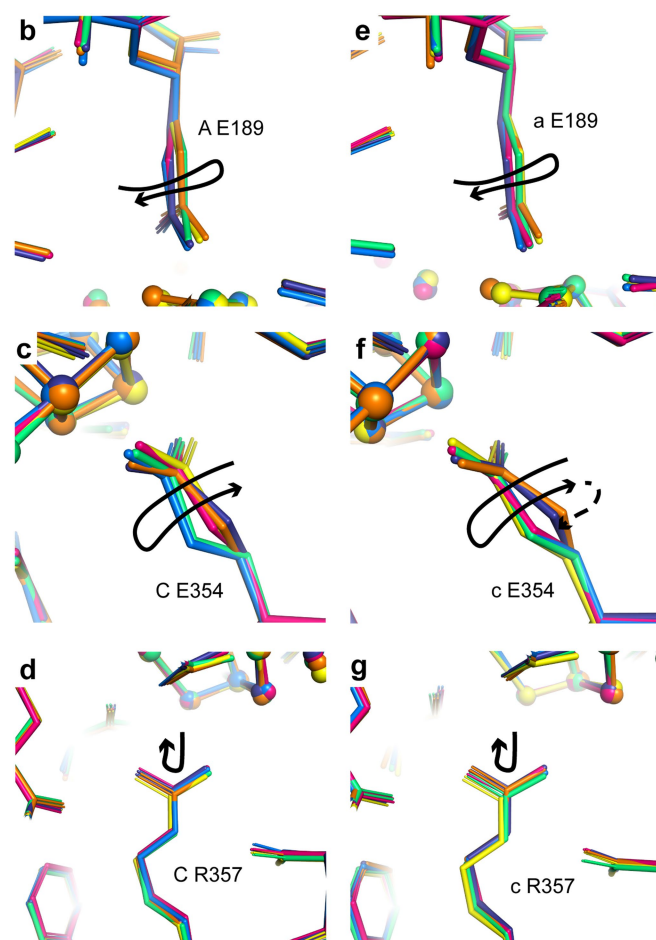
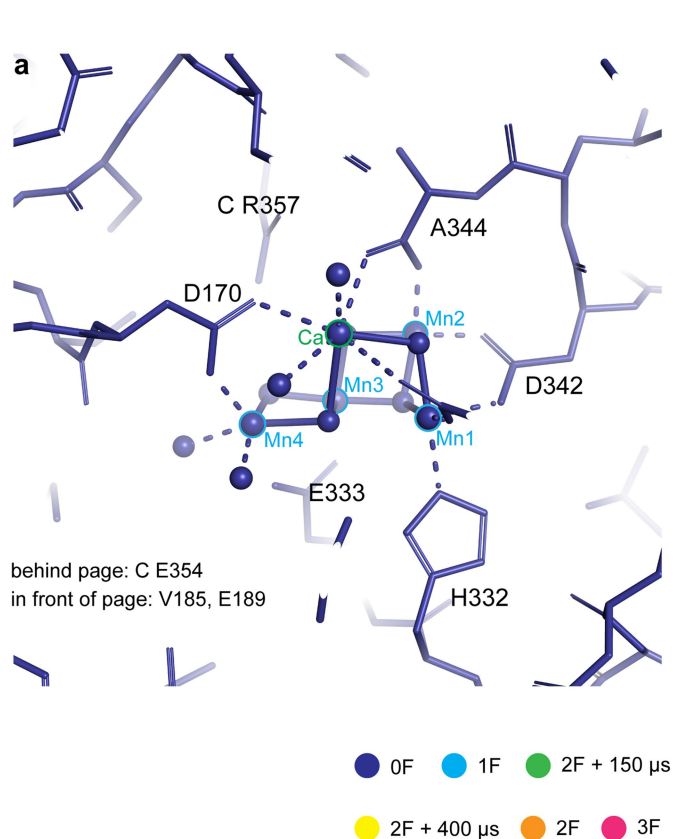
of both O<sub>2</sub> and XES data using an equal miss parameter of 22% and 100% S<sub>1</sub> population in the 0F sample (black traces in **a**, **b**; S-state distribution listed in the columns headed O<sub>2</sub>), and a direct fit of the XES data using a 8% miss parameter in the S<sub>1</sub>→S<sub>2</sub> and a 27% miss parameter for the S<sub>2</sub>→S<sub>3</sub> and S<sub>3</sub>→S<sub>0</sub> transitions (XES in **c**). For the XES fit, shifts of −0.06 eV per oxidation state increase for all S states were assumed. The XES raw spectra are published elsewhere<sup>36</sup>.





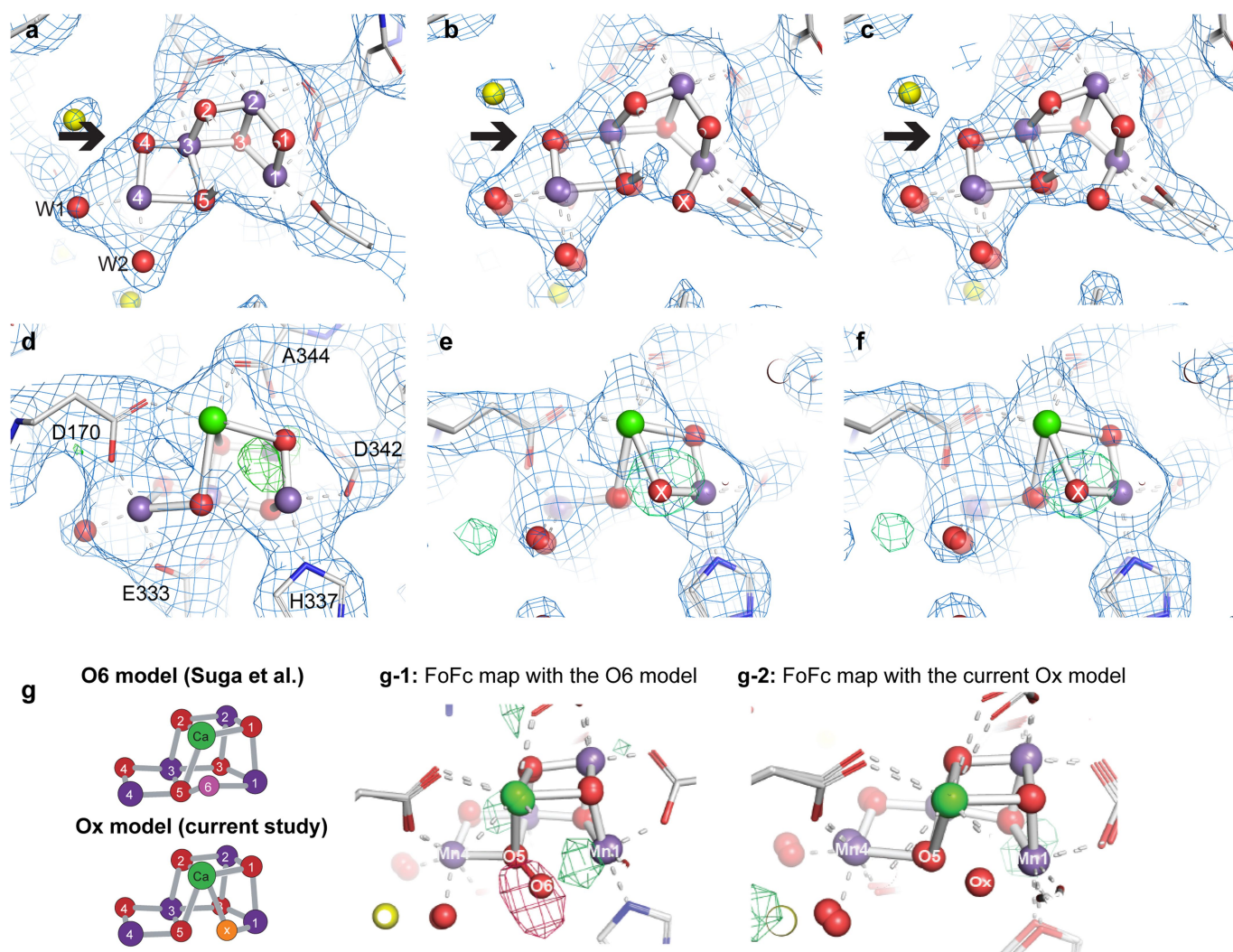
**Extended Data Fig. 3 | Isomorphous difference maps in the second monomer at the  $Q_B$  site.** **a–c**,  $F_{\text{obs}} - F_{\text{obs}}$  maps contoured at  $3\sigma$  at plastoquinone  $Q_B$  in monomer **a**. **a**, 1F – 0F difference map matching reduction of the plastoquinone to a semiquinone and concomitant slight geometry change. **b**, 2F – 0F difference map matching replacement of the

fully reduced quinol with another quinone at the original position. **c**, 3F – 0F difference map, showing again structural changes similar to the 1F – 0F map, indicating formation of the semiquinone. Similar views are shown for monomer **A** in Fig. 1d–f and comparison of both monomers indicates similar flash-induced changes in both monomers.



**Extended Data Fig. 4 | Movement of ligands around the OEC in the different S states.** **a**, Overview of the ligand environment of the OEC, showing the dark state (0F) structure. Coordination of the OEC by nearby side chains and water molecules is indicated by dashed lines. **b–g**, Trends for selected individual side chains in both monomers (**b–d**, monomer A; **e–g**, monomer a). Overlays of the refined models at the OEC following

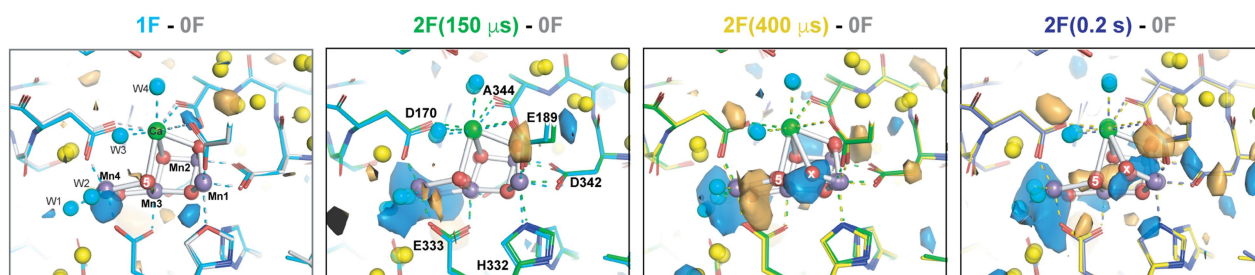
least-squares fitting of subunit D1 residues 55–65, 160–190 and 328, subunit CP43 residues 328 and 354–358, and chain D residue 352 of each other model to the 0F model. The largest and most consistent motions of side chains near the OEC through the sequence of illuminated state models are annotated with arrows indicating the trend. A motion observed in only one monomer is indicated by a dashed line.



**Extended Data Fig. 5 | Impact of the data quality on the resolving power of the maps.** **a–f**, The data quality evidenced by 2F state models and  $2mF_{\text{obs}} - DF_{\text{calc}}$  maps contoured at  $1.5\sigma$ . **a**, 5TIS (2F, 2.25 Å) model and map. Overlays indicate atom numbering in the OEC and the identities of selected coordinating sidechains. **b**, Current 2F model and map cut to 2.25 Å. **c**, Current 2F model and map at the full 2.07 Å resolution. Emergence of locations of O4 with improved data quality is indicated by bold arrow. **d**, As in **a** from a different angle and with  $mF_{\text{obs}} - DF_{\text{calc}}$  density at  $3\sigma$  indicating the lack of sufficient evidence for inserting an additional O atom at a chemically reasonable position. **e, f**, As in **b, c** from the same direction as **d** and with  $mF_{\text{obs}} - DF_{\text{calc}}$  density at  $3\sigma$  shown to 2.25 and 2.05 Å, respectively, after omitting the inserted Ox

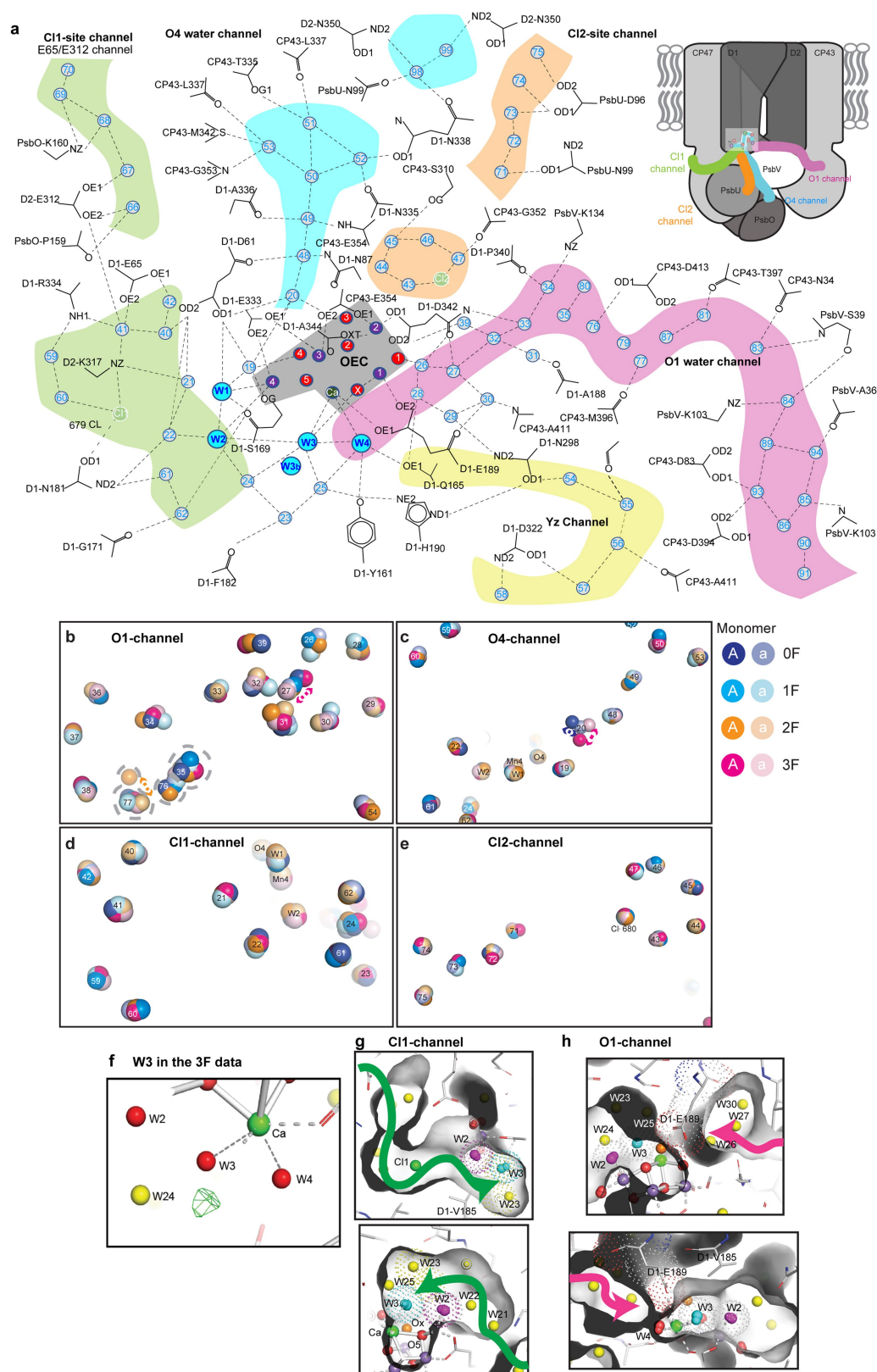
atom. Centring of the refined Ox position within the omit density gives a clear indication of the position of the inserted water in the  $S_3$  state with the current, higher-quality data, even when artificially cut to the same resolution as the previous data set. **g**,  $mF_{\text{obs}} - DF_{\text{calc}}$  maps of the 2F data that compare the O6 model from Suga et al.<sup>9</sup> and the Ox model from the current study. Map shows the  $mF_{\text{obs}} - DF_{\text{calc}}$  density calculated with our current 2F data and our model adding the O6 position of Suga et al.<sup>9</sup> (with the occupancy of 0.7 and  $B$ -factor of 30) (**g-1**), and with our Ox model (**g-2**). We see clearly a positive density for the missing Ox and a negative density at the O6 position in **g-1**. Schematics of the O6 and Ox  $S_3$  models are shown on the left.





**Extended Data Fig. 6 | Isomorphous difference maps in the second monomer at the OEC.** Isomorphous difference density OEC sites in monomer a.  $F_{\text{obs}} - F_{\text{obs}}$  difference densities between the various illuminated states and the 0F data are contoured at  $+3\sigma$  (blue) and

$-3\sigma$  (orange). The model for the 0F data is shown in light grey whereas carbons are coloured as follows: 1F (cyan), 2F (150  $\mu\text{s}$ ) (green), 2F (400  $\mu\text{s}$ ) (yellow) and 2F (0.2 s) (blue).



Extended Data Fig. 7 | See next page for caption.

**Extended Data Fig. 7 | Water environment of the OEC.** **a**, Extended schematic of the hydrogen bonding network connecting the OEC to the solvent-exposed surface of PSII and identification of several channels for either possible water movement or proton transfer. Top right, locations of four selected channels in the PSII monomer. **b–e**, Movements within the water networks across monomers. Coloured spheres are shown for each ordered water or chloride ion across the four metastable states, 0F through 3F, and for both monomers, with the stronger colour matching the first (A) monomer and the lighter colour matching the second (a) monomer. For ordered solvent, residue number is shown; for OEC atoms, the atom identifier is shown; and for the Cl2 site, the  $\text{Cl}^-$  680 label is shown. **b**, The O1 water chain. Positional disagreement between monomers is visible especially near waters 77 (2F) and 27 (3F) and is on the same scale as changes between illuminated states, both of which may indicate a more dynamic water channel. **c**, The O4 water chain. With the notable exception of water 20, most water positions are stable across monomers

and illuminated states. Water 20 is highly unstable in position in the two states (0F, 3F) in which it is modelled, and there is not sufficient density in the remaining states to model a water 20 position. **d**, The Cl1 site water channel with no notable movements. **e**, The Cl2 site water channel with no notable movements. **f**, Indication of a split position of W3 in the  $S_0$  state.  $mF_{\text{obs}} - DF_{\text{calc}}$  difference density (green mesh) in the 3F state suggests an alternate position near W3 (W3b in Fig. 4d). **g**, **h**, Possible access to W3/Ox side from the Cl1 or the O1 channel. The surface of the protein is shown in grey to visualize the extent of the cavities around the OEC, and Van der Waals radii are indicated for selected residues or atoms by dotted spheres. Shown are two different views for each channel. The direction of the Cl1 channel is indicated by a green arrow and the O1 channel by a pink arrow. Water W2 is shown in purple, W3 in cyan and Ox in orange. Yellow spheres indicate other waters. Mn are shown in magenta, other bridging oxygens as red spheres.



**Extended Data Table 1 | Merging and refinement statistics for (a) the refined data sets including all lattices or (b) for the additional data sets containing only lattices with unit cells within 1% of the target unit cell**

a	0F	1F	2F (150 $\mu$ s)	2F (400 $\mu$ s)	2F <sup>†</sup>	3F <sup>†</sup>
Resolution range refined (Å)	30.552 - 2.05	30.427 - 2.08	30.783 - 2.5	30.851 - 2.2	30.578 - 2.07	31.005 - 2.04
Resolution range upper bin (Å)	2.085 - 2.05	2.116 - 2.08	2.543 - 2.5	2.238 - 2.2	2.106 - 2.07	2.075 - 2.04
Wavelength (Å)	1.303	1.303	1.301	1.301	1.303	1.303
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell parameters (Å)	a=116.9 b=221.4 c=308.7	a=116.8 b=220.9 c=307.0	a=117.6 b=222.7 c=309.1	a=117.7 b=222.6 c=308.5	a=117.0 b=221.5 c=308.3	a=116.7 b=221.2 c=307.6
Images merged	30366	23744	4231	13072	24481	25134
Unique reflections (upper bin)	508701 (25202)	487161 (24128)	281837 (13902)	412258 (20410)	494189 (24482)	516201 (25594)
Completeness (upper bin)	99.95% (99.81%)	99.95% (99.78%)	99.92% (99.96%)	99.94% (99.91%)	99.95% (99.87%)	99.95% (99.86%)
CC <sub>1/2</sub> (upper bin)	98.7% (1.8%)	97.5% (1.3%)	93.8% (6.8%)	96.0% (0.6%)	98.2% (2.1%)	98.6% (0.8%)
Multiplicity (upper bin)	186.8 (9.0)	160.1 (8.4)	45.4 (9.2)	98.4 (9.8)	156.8 (9.6)	170.7 (9.3)
Pred. multiplicity* (upper bin)	458.3 (360.4)	383.1 (301.8)	88.2 (66.6)	272.3 (209.9)	375.1 (291.9)	413.0 (318.6)
$I/\sigma_{H_{14}}(I)^{\ddagger}$ (upper bin)	16.6 (0.5)	15.0 (0.5)	11.1 (1.0)	10.6 (0.7)	15.5 (0.6)	17.0 (0.6)
Wilson B-factor	26.8	27.1	35.5	27.4	27.3	27.1
R-factor	18.48%	18.90%	16.69%	19.33%	18.44%	18.64%
R-free	23.99%	24.56%	24.60%	26.39%	24.75%	24.85%
Number of atoms	103732	103728	103713	103719	105764	105761
Number non-hydrogen atoms	52203	52199	52188	52194	53286	53283
Ligands	186	186	186	186	188	188
Waters	2021	2017	2012	2016	2015	2010
Protein residues	5306	5306	5306	5306	5306	5306
RMS (bonds)	0.014	0.015	0.015	0.016	0.014	0.014
RMS (angles)	1.51	1.53	1.57	1.61	1.52	1.50
Ramachandran favored	96.5%	95.8%	94.8%	95.3%	96.2%	96.3%
Ramachandran outliers	0.31%	0.38%	0.56%	0.38%	0.34%	0.31%
Clashscore	5.4	6.5	6.8	6.3	6.6	6.8
Average B-factor	42.8	44.4	50.3	46.5	43.4	42.3

b	0F	1F	2F (150 $\mu$ s)	2F (400 $\mu$ s)	2F <sup>†</sup>	3F <sup>†</sup>
Resolution range refined (Å)	30.05 - 2.07	30.04 - 2.13	29.70 - 2.60	29.85 - 2.30	30.18 - 2.09	30.05 - 2.05
Resolution range upper bin (Å)	2.11 - 2.07	2.17 - 2.13	2.65 - 2.60	2.34 - 2.30	2.13 - 2.09	2.09 - 2.05
Wavelength (Å)	1.301	1.301	1.301	1.301	1.301	1.301
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell parameters (Å)	a=117.6 b=222.8 c=309.7	a=117.6 b=222.8 c=309.7	a=117.6 b=222.8 c=309.7	a=117.6 b=222.8 c=309.7	a=117.6 b=222.8 c=309.7	a=117.6 b=222.8 c=309.7
Images merged	15669	10757	2087	4576	13603	14019
Unique reflections (upper bin)	490658 (24364)	450643 (22354)	249070 (12358)	358519 (17699)	476797 (23606)	505050 (25047)
Completeness (upper bin)	99.96% (99.98%)	99.95% (99.96%)	99.91% (99.93%)	99.93% (99.93%)	99.96% (99.96%)	99.95% (99.92%)
CC <sub>1/2</sub> (upper bin)	98.2% (2.1%)	95.6% (2.8%)	92.4% (6.8%)	96.1% (2.9%)	97.7% (0.9%)	98.0% (0.4%)
Multiplicity (upper bin)	98.5 (11.1)	71.7 (10.6)	28.3 (9.8)	48.3 (10.4)	89.9 (11.0)	101.2 (10.6)
Pred. multiplicity* (upper bin)	204.8 (153.1)	144.2 (107.2)	41.3 (29.81)	86.3 (62.32)	187.6 (139.22)	210.1 (154.24)
$I/\sigma_{H_{14}}(I)^{\ddagger}$ (upper bin)	13.3 (0.7)	11.1 (0.7)	11.5 (1.1)	10.6 (0.8)	12.7 (0.7)	14.3 (0.6)
Wilson B-factor	24.4	24.4	31.8	24.0	24.8	24.9
R-factor	19.18%	19.75%	17.99%	19.82%	19.48%	19.44%
R-free	23.20%	24.43%	23.97%	24.94%	23.93%	23.62%
Number of atoms	52325	52154	50943	51500	52251	52319
Number non-hydrogen atoms	52325	52154	50943	51500	52251	52319
Ligands	187	187	187	187	187	187
Waters	2226	2054	843	1400	2151	2219
Protein residues	5300	5300	5300	5300	5300	5300
RMS (bonds)	0.008	0.009	0.009	0.009	0.008	0.008
RMS (angles)	0.98	1.06	1.08	1.07	0.97	0.97
Ramachandran favored	97.3%	97.1%	96.5%	96.8%	97.6%	97.5%
Ramachandran outliers	0.36%	0.31%	0.38%	0.27%	0.35%	0.27%
Clashscore	5.7	6.7	7.7	7.3	6.0	5.9
Average B-factor	37.6	39.9	44.2	44.4	38.9	38.6

\*Predictions multiplicity is the multiplicity of all spot predictions matching the indexing solution on a given image, before a per-image resolution cutoff. Multiplicities for data sets merged by the Monte Carlo method (for example, Kupitz et al.<sup>65</sup> and Suga et al.<sup>3</sup>) without per-image resolution cutoffs are best compared with this metric.

<sup>†</sup> $I/\sigma(I)$  calculation as described<sup>66,64</sup>.

<sup>‡</sup>For the 2F and 3F data sets, a region of 66 amino acids around the OEC was modelled as double conformers to reflect the contribution from the two main S-state species in each of the data sets with contributions of 30% and 70% for S<sub>2</sub> and S<sub>3</sub> in the 2F and 40% and 60% for S<sub>3</sub> and S<sub>0</sub>, respectively, in the 3F data sets.

**Extended Data Table 2 | Interatomic distances at the OEC in each merged data set, in each monomer (A/a), in Å**

Distances (Å)		0F		1F		2F (150 $\mu$ s)		2F (400 $\mu$ s)		2F OEI (S <sub>3</sub> )		3F OEC (S <sub>0</sub> )	
		A	a	A	a	A	a	A	a	A	a	A	a
CA1-	MN1	3.41	3.45	3.46	3.38	3.49	3.45	3.47	3.58	3.37	3.37	3.44	3.31
	MN2	3.38	3.38	3.40	3.42	3.40	3.44	3.41	3.39	3.27	3.38	3.45	3.41
	MN3	3.52	3.49	3.54	3.50	3.42	3.53	3.49	3.57	3.52	3.61	3.55	3.52
	MN4	3.77	3.88	3.90	3.90	3.72	3.96	3.84	4.07	3.90	4.11	3.94	4.07
MN1-	MN2	2.77	2.77	2.85	2.77	2.89	2.78	2.86	2.72	2.77	2.73	2.81	2.71
	MN3	3.22	3.26	3.29	3.22	3.38	3.32	3.39	3.35	3.34	3.32	3.30	3.24
	MN4	4.80	4.91	4.84	4.88	5.01	5.05	5.05	5.16	5.01	5.11	4.97	4.98
MN2-	MN3	2.87	2.83	2.86	2.82	2.85	2.83	2.88	2.82	2.85	2.86	2.90	2.82
MN3-	MN4	2.70	2.77	2.70	2.79	2.70	2.84	2.74	2.90	2.70	2.84	2.79	2.91
MN1-	Ox							1.82	1.67	1.78	1.80		
MN4-	O5							2.23	2.28	2.17	2.27		
O5-	Ox							1.84*	1.87*	2.10	2.07		
MN4-	W1	2.18	2.13	2.06	2.17	2.03	1.97	2.37	2.23	2.16	2.10	2.04	2.01
	W2	2.12	2.14	2.13	2.19	2.09	2.37	2.20	2.43	2.11	2.13	2.22	2.22
CA1-	W3	2.53	2.53	2.58	2.63	2.58	2.59	2.52	2.57	2.51	2.58	2.61	2.56
	W3B/C											3.25	4.26
	W4	2.36	2.30	2.37	2.22	2.57	2.37	2.37	2.43	2.30	2.30	2.20	2.29

\*The O5–Ox distance in the 2F (400  $\mu$ s) data set is shorter than in the 2F data set as for this data set we did not model two configurations for the OEC. Hence, the O5 position that is used to measure the O5–Ox distance is definitely influenced by the contribution of O5 in the S<sub>2</sub> state (closer to Mn1) and only partly by O5 in the S<sub>3</sub> state (longer Mn1–O5 distance).

**Extended Data Table 3 | Channel nomenclature in the literature**

This work	Ho and Styring <sup>66</sup>	Vassiliev <sup>67</sup>	Murray <sup>68</sup>	Gabdulkhakov <sup>69</sup>	Umena <sup>70</sup>	Sakashita <sup>71</sup>
01 channel	large channel	4.A	<i>channel ii</i>	B1		01-water chain
04 channel	narrow channel	2		E, F		04-water chain
CI1 channel	broad channel	1	<i>channel iii</i>	D	4.b	E65/E312 channel
CI2 channel					4.c	

Summary of the correspondence of the multiple names used for identifying the water and proton channels in PSI<sup>66–71</sup>.



## Reporting Summary

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

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

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

Our web collection on [statistics for biologists](#) may be useful.

### Software and code

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

#### Data collection

We used the live version of psana (ana-current) provided by SLAC and development versions of open source programs dials.stills\_process, the cctbx.xfel GUI and cxi.merge, which are distributed with DIALS packages publicly available at <http://dials.github.io>, for data collection. The version of the DIALS package used was 1.dev.2679-g85a721f40

#### Data analysis

We used development versions of open source programs dials.stills\_process, the cctbx.xfel GUI and cxi.merge, which are distributed with DIALS packages publicly available at <http://dials.github.io> and END/RAPID structure factor modification software available at <http://bl831.als.lbl.gov/END/RAPID/>. We also used development versions of Phenix structure refinement software available at <https://phenix-online.org> (version 1.14rc2-3192), Coot structure visualization and modification software version 0.8.8-pre, and PyMol structure visualization and figure rendering software 2.0.6.

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

## Data

Policy information about [availability of data](#)

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

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

Molecular model coordinates and crystallographic structure factors are available in the Protein Data Bank for PS II structures 0F (ID 6DHE), 1F (ID 6DHF), 2F (150  $\mu$ s) (ID 6DHG), 2F (400  $\mu$ s) (ID 6DHH), 2F (ID 6DHO) and 3F (ID 6DHP).

## Field-specific reporting

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

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

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

## Life sciences study design

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

Sample size	Sample size is stated for all data sets in the Method section and Extended Data Table 1
Data exclusions	Lattices not extending to 3 Ångstrom (or 6 Ångstrom for 2F (150 $\mu$ s) and 2F (400 $\mu$ s) datasets) resolution were excluded. Images were also excluded based on deviation of the XES signal from the expected signal for the specific illumination condition.
Replication	Measurements were conducted in two independent experiments and resulting data first evaluated separately and then after similar observations found in each individual set data from both experiments was merged.
Randomization	Different Illumination states were interleaved as much as possible during data collection within the constraints of the XFEL experiment.
Blinding	Investigators were not blinded to group allocation during data collection. Blinding was not relevant to this study as data for all different states studied was collected interleaved, ensuring similar experimental conditions. Subsequent data treatment was also performed in the same way for all data sets.

## Reporting for specific materials, systems and methods

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

### Methods

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

## Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials All materials are readily available from authors upon request.