

ARTICLE

Received 20 Feb 2014 | Accepted 10 Jun 2014 | Published 9 Jul 2014

DOI: 10.1038/ncomms5371

# Taking snapshots of photosynthetic water oxidation using femtosecond X-ray diffraction and spectroscopy

Jan Kern<sup>1,2</sup>, Rosalie Tran<sup>1</sup>, Roberto Alonso-Mori<sup>2</sup>, Sergey Koroidov<sup>3</sup>, Nathaniel Echols<sup>1</sup>, Johan Hattne<sup>1</sup>, Mohamed Ibrahim<sup>4,5</sup>, Sheraz Gul<sup>1</sup>, Hartawan Laksmono<sup>6</sup>, Raymond G. Sierra<sup>6</sup>, Richard J. Gildea<sup>1,†</sup>, Guangye Han<sup>1</sup>, Julia Hellmich<sup>4,5</sup>, Benedikt Lassalle-Kaiser<sup>1,†</sup>, Ruchira Chatterjee<sup>1</sup>, Aaron S. Brewster<sup>1</sup>, Claudiu A. Stan<sup>6</sup>, Carina Glöckner<sup>5</sup>, Alyssa Lampe<sup>1</sup>, Dörte DiFiore<sup>5</sup>, Despina Milathianaki<sup>2</sup>, Alan R. Fry<sup>2</sup>, M. Marvin Seibert<sup>2,†</sup>, Jason E. Koglin<sup>2</sup>, Erik Gallo<sup>7</sup>, Jens Uhlig<sup>7</sup>, Dimosthenis Sokaras<sup>8</sup>, Tsu-Chien Weng<sup>8</sup>, Petrus H. Zwart<sup>1</sup>, David E. Skinner<sup>9</sup>, Michael J. Bogan<sup>2,6</sup>, Marc Messerschmidt<sup>2</sup>, Pieter Glatzel<sup>7</sup>, Garth J. Williams<sup>2</sup>, Sébastien Boutet<sup>2</sup>, Paul D. Adams<sup>1</sup>, Athina Zouni<sup>4,5</sup>, Johannes Messinger<sup>3</sup>, Nicholas K. Sauter<sup>1</sup>, Uwe Bergmann<sup>2</sup>, Junko Yano<sup>1</sup> & Vittal K. Yachandra<sup>1</sup>

The dioxygen we breathe is formed by light-induced oxidation of water in photosystem II. O<sub>2</sub> formation takes place at a catalytic manganese cluster within milliseconds after the photosystem II reaction centre is excited by three single-turnover flashes. Here we present combined X-ray emission spectra and diffraction data of 2-flash (2F) and 3-flash (3F) photosystem II samples, and of a transient 3F' state (250 μs after the third flash), collected under functional conditions using an X-ray free electron laser. The spectra show that the initial O–O bond formation, coupled to Mn reduction, does not yet occur within 250 μs after the third flash. Diffraction data of all states studied exhibit an anomalous scattering signal from Mn but show no significant structural changes at the present resolution of 4.5 Å. This study represents the initial frames in a molecular movie of the structural changes during the catalytic reaction in photosystem II.

<sup>1</sup>Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. <sup>2</sup>LCLS, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. <sup>3</sup>Institutionen för Kemi, Kemiskt Biologiskt Centrum, Umeå Universitet, 90187 Umeå, Sweden. <sup>4</sup>Institut für Biologie, Humboldt-Universität zu Berlin, D-10099 Berlin, Germany. <sup>5</sup>Max-Volmer-Laboratorium für Biophysikalische Chemie, Technische Universität, D-10623 Berlin, Germany. <sup>6</sup>PULSE Institute, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. <sup>7</sup>European Synchrotron Radiation Facility, F-38043 Grenoble, France. <sup>8</sup>SSRL, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. <sup>9</sup>National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. † Present addresses: Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, UK (R.J.G.); Synchrotron SOLEIL, F-91192 Gif-Sur-Yvette, France (B.L.-K.); Department of Cell and Molecular Biology, Uppsala Universitet, 751 24 Uppsala, Sweden (M.M.S.). Correspondence and requests for materials should be addressed to U.B. (email: bergmann@slac.stanford.edu) or to J.Y. (email: jyano@lbl.gov) or to V.K.Y. (email: vkyachandra@lbl.gov).

**A**erobic life on earth is supported by the constant regeneration of dioxygen through photosynthetic water oxidation in green plants, algae and cyanobacteria. This reaction takes place in photosystem II (PS II), a multi-subunit membrane protein complex. PS II couples the one-electron photochemistry of the primary charge separation at the reaction center with the four-electron redox chemistry of water oxidation at the  $\text{Mn}_4\text{O}_5\text{Ca}$  cluster of the oxygen evolving complex (OEC) at the luminal side of the protein complex, using the spatial and temporal organization of the electron donor and acceptor cofactors in PS II<sup>1,2</sup>. This well-controlled electron and proton flow results in the high quantum efficiency of PS II.

During the water oxidation reaction, the OEC functions as a redox capacitor by storing four oxidizing equivalents before the release of molecular oxygen. Starting from the dark stable  $S_1$  state, the oxidation state of the OEC is increased by one on each light excitation of PS II until the highest oxidized stable intermediate state,  $S_3$ , is reached. Following the next light excitation, the OEC is oxidized one more time to form the transient  $S_3Y_Z^{\text{ox}}$  and  $S_4$  states that lead to dioxygen formation, which converts the OEC to its most reduced state,  $S_0$  (ref. 3). The fourth light excitation sets the OEC back to the  $S_1$  state, and thereby completes the cycle (Fig. 1a).

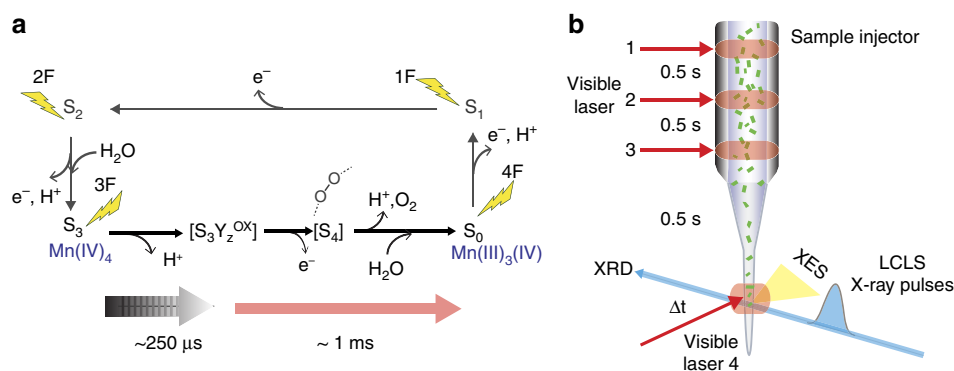
Much structural and mechanistic information about PS II, the OEC and the O–O bond formation was gained through mass spectrometric<sup>4</sup>, various spectroscopic<sup>4–9</sup>, crystallographic<sup>10–13</sup> and theoretical<sup>14–16</sup> studies over the past decade. In particular, the most recent structure, inferred from X-ray diffraction (XRD) data, has provided detailed geometric information of the OEC, including ligands and bound water molecules<sup>13</sup>. Most of the experimental studies, however, are carried out at cryogenic temperatures and represent a static picture of the system in a frozen state. Although the stable intermediate states,  $S_0$  through  $S_3$ , can be trapped and studied at cryogenic temperatures, the critical  $S_3 \rightarrow S_3Y_Z^{\text{ox}} \rightarrow S_4 \rightarrow S_0$  step—where dioxygen is formed, two protons and  $\text{O}_2$  are released, and where at least one substrate water binds—only occurs under ambient conditions and has no intermediates that can be cryotrapped. To date, there has been only one transient X-ray spectroscopy study of the  $S_3 \rightarrow S_3Y_Z^{\text{ox}} \rightarrow S_4 \rightarrow S_0$  transition performed at room temperature (RT)<sup>17</sup>. More detailed investigations of the transient states by X-ray spectroscopy and by kinetic crystallography have been hampered due to the severe radiation damage, especially to the  $\text{Mn}_4\text{CaO}_5$  cluster that is significantly faster at RT as compared with cryogenic conditions. However, X-ray-induced changes, particularly at the redox-active metal site, have even been an issue for experiments carried out at cryogenic temperatures<sup>18–20</sup>.

We have recently introduced a combined spectroscopy and diffraction data collection methodology at RT<sup>21</sup> using the ‘probe before destroy’ method<sup>22–24</sup> made possible by the ultra-short (fs) and bright X-ray pulses of an X-ray free electron laser (XFEL). In this approach, XRD data and Mn  $K\beta$ -X-ray emission spectra (XES), sensitive to the metal charge density<sup>25,26</sup>, are measured simultaneously from micrometre-sized crystals of PS II, thereby obtaining information about the geometric and the electronic structure of the active site, under identical conditions. Owing to the ultra-short fs X-ray pulse duration, the sample is probed before the manifestation of X-ray induced changes—which predominantly take place on the picosecond time scale (for damage to the atomic structure)—even under ambient conditions. One should note that with conventional synchrotron X-ray sources the main source of radiation damage is via the generation of radicals from the solvent (water). Subsequent diffusion of these radicals leads to specific damage (for example, reduction of metal sites) and modification of amino acid side chains (for example, decarboxylations). Such events are diffusion controlled and occur on a longer time scale (> picoseconds) and seem not to be dependent on the dose rate. Earlier work<sup>21,27</sup> showed that the approach of using ultrafast (<50 fs) and ultra-bright ( $10^{12}$  photons per pulse) X-ray pulses permits the collection of XES and XRD data from intact PS II, and we reported results from the dark-adapted ( $S_1$ ) and the one flash ( $S_2$ ) samples with an XRD resolution limited to 5.5 Å.

Here we present XES and XRD data from the last step of the Kok cycle, where  $\text{O}_2$  is evolved, with an improved resolution of 4.5 Å. This step, triggered by the third flash given to dark-adapted PS II samples, advances the PS II complex from the  $S_3$  to the  $S_0$  state, via the transient  $S_3Y_Z^{\text{ox}}$  and  $S_4$  states ( $S_3 \rightarrow S_3Y_Z^{\text{ox}} \rightarrow S_4 \rightarrow S_0$  transition). Furthermore, we observe an anomalous signal for the Mn atoms in the OEC from all the states, including the transient  $S_3Y_Z^{\text{ox}}$  state. This observation supports the quality of our XRD data and also the data analysis protocols, and we envision that the Mn anomalous signal could be used as a sensitive probe for monitoring changes of the atomic positions of Mn in the OEC during the catalytic cycle in future studies at higher resolution.

## Results

**XES at different time points in the catalytic cycle.** PS II was advanced through its reaction cycle *in situ*, using a flow/illumination scheme (Fig. 1b) employing an electrospun liquid jet<sup>28</sup>. The protocol consisted of visible-laser illumination using three optical fibres directly attached to the sample delivery capillary, and an additional laser for illumination of the sample in the jet



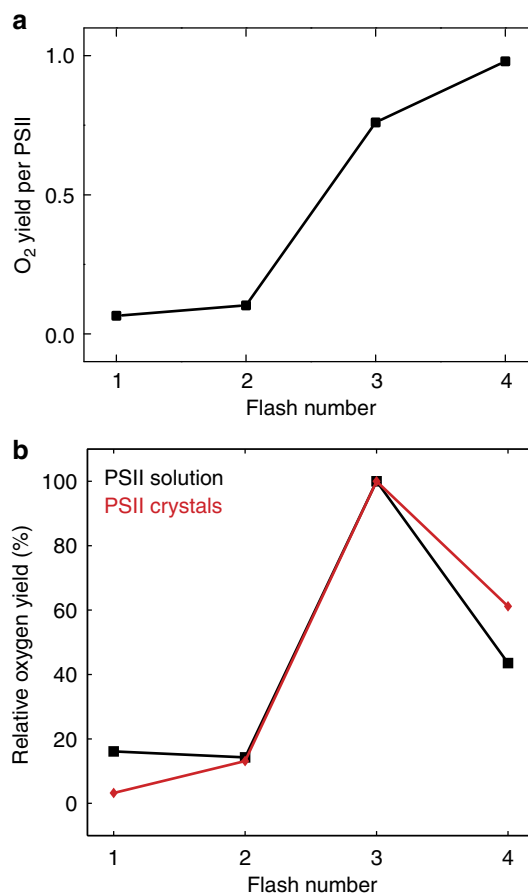
**Figure 1 | Flash-induced changes in PS II and experimental set-up used at LCLS. (a)** Kok cycle describing the different stable intermediate states of the catalytic water oxidation reaction in PS II. **(b)** Scheme for the illumination set-up used to advance PS II in the catalytic cycle and measure simultaneously the XRD and XES signal at LCLS. Lasers 2 and 3 were used to generate 2F samples, lasers 1, 2 and 3 for 3F samples and lasers 2, 3 and 4 to generate the 3F' samples.

(see Methods). The temporal frequency for illumination was chosen to match the sample flow rate, so that each volume segment was illuminated by each fibre once while passing through the capillary. The set-up also allows enough time ( $\sim 0.5$  s) for complete PS II turnover between consecutive illuminations, which takes into account the slower acceptor side reactions<sup>7,29–31</sup>, while being rapid enough to avoid significant decay of the  $S$ -states, that are stable for on the order of several tenths of seconds. The fourth laser (labelled ‘laser 4’ in Fig. 1b) illuminated the sample in the jet, to study transients during the  $S_3$  to  $S_0$  transition by changing the timing between the third visible-laser pulse and the X-ray probe pulse.

$O_2$  detection via membrane-inlet mass spectrometry (MIMS) was used for optimizing the conditions for  $S$ -state turnover in the capillary flow sample delivery system, using a facsimile of the flow/illumination set-up employed at Linac Coherent Light Source (LCLS) (see Methods). One of the most important factors in the illumination scheme is the required light intensity for efficient turnover through the  $S_i$  state cycle. Too low intensities can lead to only partial turnover of the samples, while too high intensities increase the miss parameter via light scattering along the capillary, and may also inactivate the sample. The optimal light intensity can be found by the quality of the  $O_2$  oscillation pattern, and also by the total  $O_2$  produced per PSII complex and flash number. The former method should normally be sufficient, but a small uncertainty remains whether there can be a certain part of the sample that never sees any light, and thus does not contribute to the oscillation pattern. To address this question, the latter method needs to be employed (see Methods), which requires the absolute calibration of the MIMS signals. The amount of  $0.73 O_2/RC$  after three flashes (Fig. 2a) shows that the light conditions used for illumination are optimal for saturating all PSII reaction centres in the sample. The  $O_2$  evolution patterns obtained from PS II solutions and PS II microcrystals (Fig. 2b) show light-induced turnover of the catalytic cycle as expected. Analysis of the flash pattern indicates that the  $S_3$  state is the majority component ( $\geq 55\%$ ) in the samples given two visible-laser flashes (2F) with virtually no  $S_0$  state present. In contrast, the largest component in the 3F samples is the  $S_0$  state ( $\geq 40\%$ ). Therefore, the difference between the 3F and 2F samples is dominated by the formation of the  $S_0$  state at the expense of the  $S_3$  state.

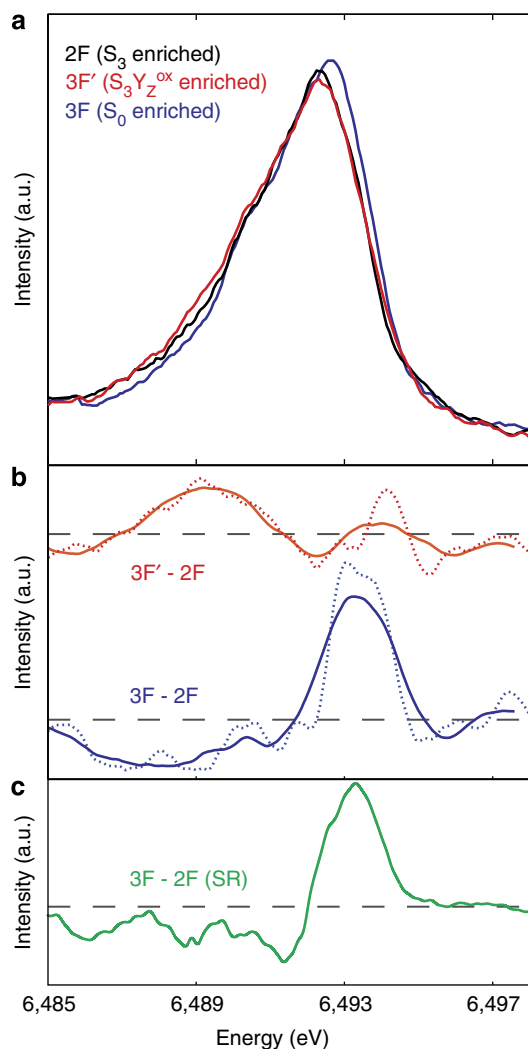
We measured XES on PS II solutions at the Coherent X-ray Imaging (CXI) instrument<sup>32</sup> at LCLS (see Methods). As shown in Fig. 3, a clear shift between the 2F ( $S_3$ -enriched) and 3F ( $S_0$ -enriched) spectra is observable. Calculation of the first moment (see Methods) revealed that the 3F spectrum is shifted  $\sim 0.1$  eV to higher energies indicating a reduction of Mn<sup>26</sup>, as expected for the transition of the OEC from the highly oxidized  $S_3$  to the most reduced  $S_0$  state, in which the formal oxidation states are assigned as Mn<sup>IV</sup><sub>4</sub> and Mn<sup>III</sup>Mn<sup>IV</sup>, respectively<sup>1,2,6,33,34</sup>. Comparing these data to synchrotron radiation (SR) data collected at cryogenic temperature from *Thermosynechococcus elongatus* PS II and previously recorded data from spinach PS II<sup>35</sup> shows a very similar trend (Fig. 3c and Supplementary Fig. 1).

In addition to the 2F and 3F spectra, we measured the XES at a time point  $250 \mu\text{s}$  after the third flash (3F') using lasers 2, 3 and 4 (Fig. 1). The XES for this transient state is similar in position to the 2F spectrum (Fig. 3a,b), but its shape is different with broadening towards the lower energy side. Although such broadening could be caused by oxidation of a fraction of the lower  $S$ -states ( $S_1$  and  $S_2$ ) in our sample, it could also be due to light-induced changes in the electronic structure of the  $S_3$  fraction. Nevertheless, the result shows that there is no significant reduction or oxidation of the Mn taking place within the  $250\text{-}\mu\text{s}$  time span between the third visible-laser excitation pulse and the X-ray probe pulse.



**Figure 2 | Oxygen production by PSII.** (a)  $O_2$  yield per PSII reaction centre as detected by MIMS as a function of flash number (measurement shown is for PS II solutions, flow rate  $0.5 \mu\text{l min}^{-1}$ , frequency 4 Hz, light intensity was  $7 \mu\text{J}$  for each fibre). (b)  $O_2$  yield measured by MIMS as a function of flash number from PS II solutions (black) and PS II microcrystals (red).

**XRD in the higher  $S$ -states.** XRD data from 2F ( $S_3$ -enriched), 3F ( $S_0$ -enriched) and 3F' ( $S_3Y_Z^{\text{ox}}$ -enriched;  $250 \mu\text{s}$  after the third flash) PS II crystals, as well as in the dark state ( $S_1$ ), were collected. Microcrystals of PS II were prepared using a new seeding protocol (see Methods). Clear Bragg spots were observed to a resolution of  $\sim 4.1 \text{ \AA}$ , with thermal diffuse scattering extending well beyond this to  $\sim 3.0 \text{ \AA}$ , indicative of correlated atomic motion in the crystal. For the 2F data, a total of 16,973 indexed patterns were merged resulting in a data set of  $4.5 \text{ \AA}$  resolution (see Table 1 and Supplementary Tables 1 and 2 for details). The resolution cutoff for the merged data sets was chosen based on the resolution dependence both of the multiplicity and of  $CC_{1/2}$ , the correlation coefficient of semi-data sets merged from odd- and even-numbered images<sup>36</sup>; that is, completeness  $> 90\%$ , multiplicity  $> 6$  and  $CC_{1/2} > 30\%$ . Likewise, data sets of 3F, 3F' and 0F states were obtained with resolutions of  $4.6 \text{ \AA}$  (13,094 lattices),  $5.2 \text{ \AA}$  (7,850 lattices) and  $4.9 \text{ \AA}$  (6,695 lattices), respectively (Table 1 and Supplementary Tables 1 and 3–5). Electron density maps for all four states are shown in Fig. 4 and Supplementary Figs 2 and 3. A comparison with the SR data cut to the same resolution shows that the level of detail visible is as expected for this resolution range (Supplementary Fig. 4). The occupancy for selected non-protein molecules was set to zero and the simulated annealing omit maps were computed for all data sets, to remove potential model bias arising from phasing with a complete, high-resolution starting model (pdb: 3bz1)<sup>12</sup>. The



**Figure 3 | Mn K $\beta$  XES of PS II.** (a) XES recorded with  $<50$  fs X-ray pulses at LCLS. Spectra were measured 0.5 s after two laser flashes (2F, black; lasers 2 and 3 on), or 0.5 s after three laser flashes (3F, blue; lasers 1, 2 and 3 on), and  $\sim 250$   $\mu$ s after three laser flashes (3F', red; lasers 2, 3 and 4 on), respectively. (b) Difference between the Mn K $\beta$  XES of PS II, blue: 3F-2F; red: 3F'-2F. Before calculating the difference curves, spectra were smoothed by moving average (dotted line) or cubic polynomial fitting (solid line, similar to the procedure used for analysing the synchrotron data). (c) The 3F-2F difference spectrum (green) from SR data collected at 15 K.

result clearly shows the electron density of the Mn<sub>4</sub>O<sub>5</sub>Ca cluster, the non-haem Fe, the chlorophyll and even partially for the quinone cofactors (Supplementary Fig. 3) in the  $mF_o - DF_c$  difference maps. The regions around the OEC, the acceptor-side quinones and non-haem iron, where the largest changes are expected, were inspected for changes between the different states. No statistically significant changes were observed in the  $2mF_o - DF_c$  maps of the individual data sets (Fig. 4a,b and Supplementary Figs 2 and 3) and in the isomorphous difference maps ( $mF_o - mF_o$ ) between the different data sets (Fig. 4c,d and Supplementary Fig. 5). This shows that any structural changes related to the S-state transitions are smaller than what we can detect at the current resolution. However, it should be noted that the  $mF_o - DF_c$  Fourier maps contain several features that are observed consistently in both monomers and all flash states; namely, an electron density peak at the position of the OEC when viewed at a contour level of  $+3\sigma$ , a small peak 10 Å distant that appears to be coordinated by residues Glu 333 and Asp 61 of the

D1 polypeptide, and other nearby peaks. Smaller negative peaks are seen at the  $-3\sigma$  contour, for example, close to Val 185 and Phe 182 of the D1 protein (Supplementary Fig. 6). In general, we observe these low-intensity peaks at the same positions in both monomers and across all four illuminated states. This suggests that they are not artefacts of the Fourier transform, and are rather due to structural differences between SR data collected at cryogenic temperature and the RT data presented here. However, the current resolution does not allow them to be fully modelled in our final atomic coordinate sets.

#### Measurement of anomalous XRD signal from Mn in PS II.

Accurate determination of the Bragg spot intensities and the derived structure factors is challenging for single-shot crystallography at XFELs<sup>21,23,24</sup>. As a control to validate the data quality and our analysis protocol, we investigated whether small anomalous differences could be detected in the recorded Bragg spot intensities. Such differences between inversion-related Bragg spots (Bijvoet pairs) arise from the collection of diffraction data at energies above an absorption edge and are often only in the order of  $\sim 1\%$  of the total signal intensity. We used an incident energy of 7.1 keV in our current XES/XRD data collection, which is close to the Mn edge (6.54 keV), and favours observing the anomalous signal from Mn in the OEC.

As a positive control of the methodology, we first analysed microcrystal diffraction data from a model system, thermolysin, which natively binds one Zn and several Ca ions<sup>37</sup>. Data from thermolysin microcrystals were collected at 1.27 Å (9.76 keV),  $\sim 100$  eV above the Zn edge (9.66 keV). Diffraction was observed out to the corners of the detector (1.50 Å) and the integrated intensities were merged to obtain a data set to 1.80 Å resolution (Table 1 and Supplementary Table 6). Analysis of the Bijvoet pairs in the merged data showed a clear anomalous signal contribution, and anomalous difference maps showed a clear maximum,  $18\sigma$  above the mean, located at the position of the Zn ion as well as lower maxima for three of the four Ca ions and for the sulphur of one of the methionine residues (Fig. 5a,b and Supplementary Fig. 7).

In PS II, a clear anomalous signal (Fig. 5c,d, Fig. 6 and Supplementary Figs 8 and 9) from Mn in the OEC is also detected in all four data sets (0F, 2F, 3F and 3F') (Supplementary Table S1). Figure 5c,d shows the anomalous difference map from the 3F data after omitting the OEC and performing simulated annealing refinement. It is evident from the overview shown in Fig. 5c that the largest peak ( $\sigma > 6$ ) in the anomalous density is located at the position of the OEC. The density covers the Mn ions in the cluster and does not include the Ca (Fig. 5d) as expected from the weaker anomalous contribution of Ca at 7.1 keV ( $f''$  of 1.6 for Ca compared with 3.4 for Mn at 7.1 keV). Similar results were obtained for the other PS II data sets for both monomers in the PS II dimer (Fig. 6 and Supplementary Figs 8 and 9). It should be noted however that the anomalous difference Patterson maps did not reveal peaks above the noise level attributable to Mn. This result is expected as also the anomalous data measured at SR sources at 3.5 Å resolution<sup>10</sup> did not yield any peaks in the Patterson map above the noise level, due to the large protein mass and the low number of anomalous scatterers per unit cell volume.

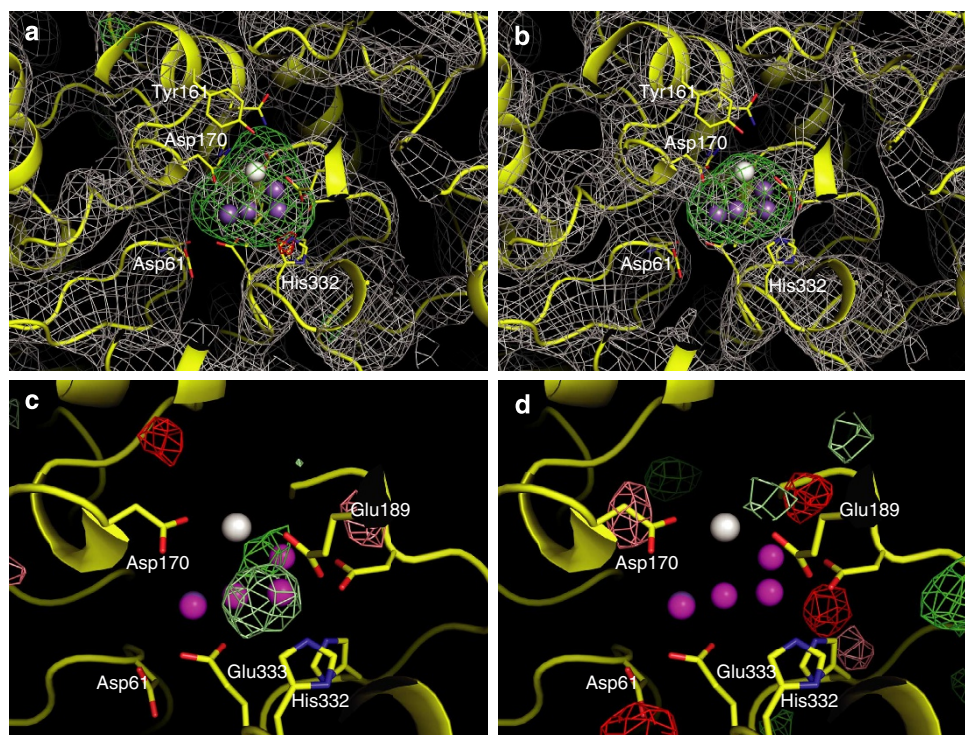
#### Discussion

The quality of the PS II XRD data reported here for the S<sub>1</sub> state is improved compared with the previously obtained XFEL data: 4.5 Å versus 5.7 Å<sup>21</sup>. Owing to the inherent fluctuations in pulse intensity, crystal size and crystal quality in single-shot microcrystal experiments at an XFEL, the signal strength varies

**Table 1 | Statistics for processed data and refined structures.**

	Dark (S <sub>1</sub> )	2-flash (2F)	3-flash + 250 μs (3F')	3-flash + 500 ms (3F)	Thermolysin
Wavelength (Å)	1.77				1.27
Resolution range (Å)	72.93–4.9 (5.08–4.9)	72.97–4.5 (4.66–4.5)	68.41–5.2 (5.39–5.2)	72.96–4.6 (4.76–4.6)	34.27–1.80 (1.86–1.80)
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 6 <sub>1</sub> 22
Unit cell dimensions (Å)	132.9 229.0 307.7	132.3 228.7 308.0	132.6 229.3 306.8	132.4 228.8 307.9	93.0 93.0 130.4
Unique reflections	41,292 (4,013)	52,965 (5,008)	34,679 (3,378)	49,771 (4,812)	31,458 (3,075)
Completeness (%)	99.7 (98.6)	99.5 (95.8)	99.7 (98.1)	99.7 (98.2)	100.0 (100.0)
Wilson B-factor	172	153	176	159	16.4
R-work	0.281 (0.363)	0.276 (0.367)	0.271 (0.347)	0.278 (0.371)	0.208 (0.349)
R-free	0.292 (0.337)	0.284 (0.393)	0.289 (0.378)	0.284 (0.346)	0.232 (0.368)
Number of non-hydrogen atoms	50,244				2,740
Macromolecules	41,052				2,415
Ligands	9,192				5
Waters	0				324
Protein residues	5,214				315
RMS (bonds) (Å)	0.005	0.005	0.005	0.005	0.005
RMS (angles) (°)	0.75	0.75	0.77	0.75	0.92
Ramachandran favoured (%)	91	91	91	91	95
Ramachandran outliers (%)	1.2	1.2	1.1	1.2	0
Clashscore	9.43	9.45	9.50	9.34	1.72
Average B-factor (Å <sup>2</sup> )	207	174	208	180	19.6

Statistics for the highest-resolution shell are shown in parentheses. All unit cell angles are 90° for photosystem II structures, and  $\alpha = \beta = 90^\circ$ ,  $\gamma = 120^\circ$  for thermolysin.

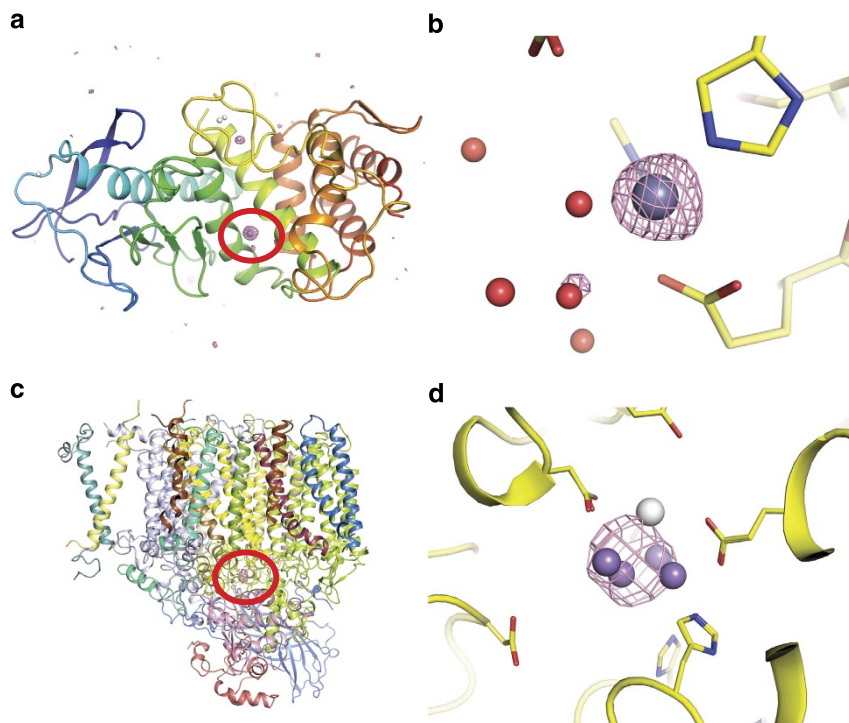


**Figure 4 | Electron density maps obtained for PS II.** (a)  $2mF_o - DF_c$  maps for the dark and (b) the 2F data of PS II are shown in grey contoured at  $1.0\sigma$ ,  $mF_o - DF_c$  maps after omitting the OEC are shown in green and red, contoured at  $\pm 5.0\sigma$ . (c)  $mF_o - mF_o$  isomorphous difference maps for the 2F-dark data and (d) the 3F-2F data are shown for both monomers and are contoured at  $+3\sigma$  (bright green, monomer I; pale green, monomer II) and  $-3\sigma$  (red, monomer I; salmon, monomer II) together with the model for the 2F data (Mn shown as magenta spheres, Ca as white sphere).

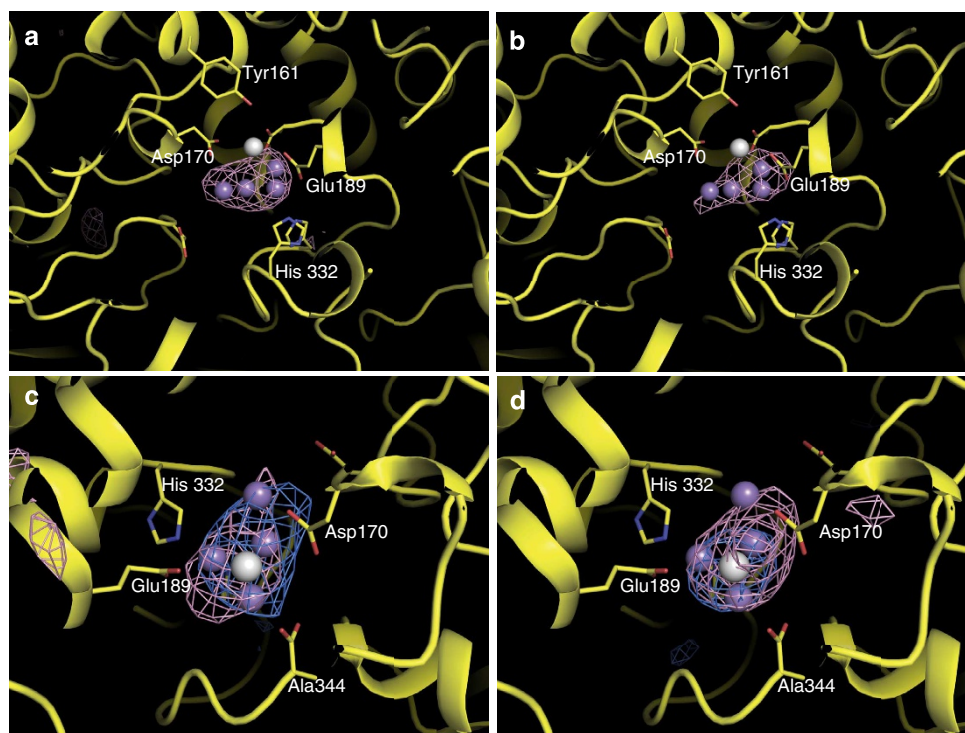
from shot to shot. Therefore, we expect a distribution of diffraction images with different maximum resolution. To avoid adding noise into the diffraction data, an individual resolution cutoff was computed for each diffraction image based on the signal strength (see Methods). The observed distribution of the resolution for PS II as well as for thermolysin explains why

the multiplicity in both cases (Supplementary Tables 2–6) decreases steadily in the higher resolution shells.

Recently, the first observation of an anomalous signal from femtosecond diffraction experiments with microcrystals at an XFEL<sup>38</sup> and the first successful *de novo* phasing of lysozyme at 2.1 Å resolution using the anomalous signal of gadolinium



**Figure 5 | Anomalous signal in the XFEL data sets.** (a) Anomalous difference map of the thermolysin data after simulated annealing with the occupancy for Zn and Ca set to zero to minimize model bias. The map is contoured at  $4.0\sigma$ , extending over the entire thermolysin molecule. The position of the highest peak in the map (Zn atom) is highlighted. (b) The same anomalous difference map of thermolysin shown in the region of the natively bound Zn ion (magenta sphere), contour level at  $3.0\sigma$ . (c) Anomalous difference map obtained from the 3F data of PS II, shown for one monomer, location of the strongest peak is highlighted, contour level at  $4.0\sigma$ . (d) Enlarged view of the 3F anomalous density for the region of the OEC (contoured at  $4.0\sigma$ ; Mn shown as magenta spheres, Ca as white sphere). All maps shown are anomalous difference simulated annealing omit maps.



**Figure 6 | Anomalous signal from Mn for different illumination states of PS II.** (a) Anomalous map of the OEC in PS II is shown for the 2F data (magenta) in monomer I. (b) Anomalous map of the 3F data in monomer I. (c) Anomalous map of the 2F (cyan) and 3F (magenta) data in monomer I, orientation is rotated by  $90^\circ$  around horizontal and vertical axis compared with the view in a. (d) Anomalous map for monomer II, 2F (cyan) and 3F (magenta) data are shown, view direction is similar to c. All maps shown are anomalous difference simulated annealing omit maps contoured at  $3\sigma$ .

obtained in an XFEL experiment were reported<sup>39</sup>. In the gadolinium phasing experiment of lysozyme, the anomalous signal strength was around 5–15%<sup>39</sup>. In comparison, we expect an anomalous signal of <1% for Mn in PS II and of ~1.5% for Zn in thermolysin. The observation of the very strong anomalous peak for Zn in the thermolysin data indicates that the Bragg peak intensities were determined with sufficient accuracy to extract the weak anomalous difference (see ref. 40 for a report on determining the anomalous Zn signal of thermolysin from SR measurements). Furthermore, the presence of the anomalous density for Mn in all of the PS II data sets, despite the expected low signal strength, confirms the quality of the data and implies that structure factors can be extracted reliably from the current PS II data sets. In this regard, it should be noted that even the high-resolution shell of the data still contains a considerable amount of anomalous signal as can be seen in Supplementary Fig. 9B.

As described above, the electron density of PS II shows the level of detail expected at the specified resolutions (4.5–5.2 Å, depending on number of collected images per S-state; Supplementary Fig. 4). The quinone co-factors ( $Q_A$  in the dark, 2F and 3F data;  $Q_B$  in the 3F data), that were not visible in the previous XFEL data due to limited resolution, are now partially visible in the  $mF_o - DF_c$  difference maps (Supplementary Fig. 3). In the earlier SR XRD structures of PS II<sup>10,41–44</sup> with a resolution lower than 3.0 Å, it was difficult to locate them with confidence (especially the mobile  $Q_B$ ) due to partial occupancy and quinone mobility.

The native XRD data indicate that there are no large-scale rearrangements of the  $Mn_4CaO_5$  cluster and its protein environment between the different states (dark, 2F, 3F and 3F') in PS II. This is in line with Mn EXAFS data, which suggests that the largest possible changes in Mn–Mn distances on S-state transitions are <0.5 Å (this estimate is based on the changes proposed in the  $S_2$  to  $S_3$  transition, if di- $\mu$ -oxo bridged Mn becomes mono- $\mu$ -oxo bridged)<sup>20</sup>, which are well below the sensitivity of our current XRD measurements. To test the level of change detectable using our data, we simulated a shift of the  $Mn_4CaO_5$  cluster by 1 and 0.4 Å compared with the starting model based on SR data. In both cases, a strong positive and negative peak corresponding to the shift was visible in the isomorphous difference density (Supplementary Fig. 10A,B), indicating that shifts of that magnitude should be detectable within the signal to noise level of the current data. In contrast, when only perturbing the position of a single Mn within the  $Mn_4CaO_5$  cluster by 0.5 Å, no clear peaks were observed in the difference density (Supplementary Fig. 10C), indicating that a structural change of such order can not be resolved at the present resolution. We furthermore evaluated the noise level in the current electron densities by computing the difference in position of the  $Mn_4CaO_5$  cluster between the two monomers in each data set. For this approach, we superimposed the  $mF_o - DF_c$  Fourier omit maps after simulated annealing (see Methods) for the two monomers and evaluated the differences between the OEC peak in these maps. The values are in the range of 0.3–0.6 Å, indicating that changes in Mn positions larger than ~0.5 Å should be visible in our data.

The height and volume of the anomalous XRD difference map peak of the OEC reflects the amount of data available for each S-state data set. Supplementary Fig. 9A shows that after scaling the anomalous difference densities of the 2F, 3F and 3F' data sets individually at a  $\sigma$ -value of 80% of the peak maxima, the extent of the anomalous peak at the OEC is roughly similar for all data sets. It is also evident from this plot that the peak position of the anomalous signal for all data sets is in the vicinity of the centre of mass of the four manganese, as expected for a signal originating from Mn. Before attempting to interpret the visible differences in

the anomalous signal for the different S-states in terms of structural differences of the  $Mn_4CaO_5$  cluster, it should be noted that at the present resolution the difference of the anomalous signal between two monomers in one PS II dimer is larger than the differences for the  $Mn_4CaO_5$  cluster in the same monomer in different S-states. We quantified this by evaluating the difference in the peak position for the OEC between the two monomers for each of the data sets and found differences in the order of above 1 Å. This indicates that the noise level in our present anomalous data precludes the observation of the small changes expected between the different illumination states. Therefore, analysis of detailed structural changes based on the anomalous signal of the OEC has to await higher data quality, and in the light of the present data interpretations of structural changes can only be made using the isomorphous omit maps, as these are derived from the full data for each flash state instead of only the anomalous data and consequently show lower noise levels.

We have used fs XES to follow the changes in the oxidation state of the  $Mn_4CaO_5$  cluster on advancing from the  $S_3$  to the  $S_0$  state. The peak shift that was observed between the 2F and the 3F sample reproduced the shift in oxidation state between the highest oxidized ( $S_3$  state) and most reduced state ( $S_0$  state) found earlier<sup>35</sup>. Interestingly, the peak position of the spectrum observed at the transient time point, 250  $\mu$ s after the third flash, is very similar to that of the 2F spectrum. The observation that there is no significant change in the oxidation state of Mn within this time span is consistent with the kinetics of Mn oxidation/reduction in the  $S_3 \rightarrow S_3 Y_Z^{Ox} \rightarrow S_4 \rightarrow S_0$  transition based on earlier time-resolved ultraviolet–visible<sup>45,46</sup>, electron paramagnetic resonance<sup>47</sup>, infrared spectroscopy<sup>7</sup> and Mn K-edge XAS<sup>17</sup> studies. It was inferred from these studies that a de-protonation step, forming a transient state  $S_3 Y_Z^{Ox}$ , (also referred to as  $S_3'$ )<sup>7,46</sup> occurs before Mn redox chemistry<sup>7,46</sup>. This lag phase before the onset of Mn redox chemistry was reported to be in the range of 100–250  $\mu$ s<sup>7,17</sup>. Our results provide direct evidence that Mn redox chemistry does not occur within the first 250  $\mu$ s after illumination of the  $S_3$  state. This implies that the formation of a  $Mn^V$  species that has been invoked for a nucleophilic attack mechanism<sup>2</sup>, or the formation of a peroxide intermediate that will result in Mn reduction<sup>1</sup>, does not occur within the first 250  $\mu$ s after the third flash. The formation of an oxygen radical species<sup>4,6</sup> within this time period cannot be excluded by our data (as no Mn oxidation would be involved), but is unlikely on the basis of the earlier ultraviolet–visible data<sup>45</sup>. The long delay before the onset of Mn redox chemistry suggests that the formation of the  $S_3 Y_Z^{Ox}$  state is not a simple deprotonation step. It is rather likely to be accompanied by slower structural changes of the  $Mn_4CaO_5$  cluster and/or the protein framework<sup>7</sup> that are required to stabilize the deprotonated  $S_3 Y_Z^{Ox}$  state of the OEC in the conformation necessary for subsequent O–O bond formation. In line with this conclusion, a recent report<sup>46</sup> underlined the importance of the exact structure of the H-bonding network for efficient turn over in the  $S_3 \rightarrow S_3 Y_Z^{Ox} \rightarrow S_4 \rightarrow S_0$  transition by demonstrating that Ca/Sr exchange in the OEC perturbs the H-bonding network and results in a significant slowing of the  $S_3 \rightarrow S_3 Y_Z^{Ox}$  transition<sup>4,46</sup>.

In summary, we have investigated the S-state intermediates and a transient state with fs XES/XRD by following the S-state transitions under ambient conditions. Advancement of S-states by *in situ* photoexcitation was confirmed by the  $O_2$  evolution pattern and the XES spectral shifts. The XES data indicate that the Mn oxidation state does not change within 250  $\mu$ s after the illumination of the  $S_3$  state. The most probable explanation for this observation is that the deprotonation process of the OEC proceeds before the electron transfer, and the O–O bond formation occurs >250  $\mu$ s after the third photo-excitation, in

agreement with previous studies<sup>7,17,46,47</sup>. Structural changes that are large enough to access with the current XRD resolution of 4.5–5.2 Å were not observed in the OEC, surrounding amino acid residues, or the quinone sides on the  $S_3$ ,  $S_3Y_2^X$  and  $S_0$  state formation, which implies that the structural changes in the OEC are within the order of  $\leq 0.5$  Å. Interestingly, our RT structural data clearly show the presence of several features—although not interpretable at the present resolution—in the  $mF_o-DF_c$  difference electron density maps, indicating structural differences in the RT XFEL data compared with the previous SR cryogenic structural models. Future improvements in the crystal quality and data, especially the anomalous signal, will also allow us to use the XFEL approach to resolve the sequence of the important structural and electronic changes during the  $S_3 \rightarrow S_3Y_2^X \rightarrow S_4 \rightarrow S_0$  transition, providing unprecedented experimental insights into the mechanism of photosynthetic water oxidation.

## Methods

**Sample preparation.** PS II was purified from *T. elongatus* as described elsewhere<sup>48</sup>. Crystals that were obtained as described in ref. 48 and a seed kit (Hampton Research, CA, USA) were used to produce a PS II seed stock solution in buffer A (100 mM PIPES, pH 7, 5 mM  $CaCl_2$ , 6% (w/w) PEG 2000, 0.03%  $\beta$ -dodecyl maltoside) for microcrystallization of PS II. Microcrystals of PS II were obtained by mixing aliquots of the PS II seed stock solution with PS II solution (chlorophyll (Chl) concentration 4 mM, corresponding to a protein concentration of  $\sim 40$  mg ml<sup>-1</sup>) in a 1:4 ratio. Box-shaped crystals (5–10  $\mu$ m in the longest dimension, 5  $\mu$ m in the shorter dimension) were suspended in buffer C (100 mM MES, pH 6.5, 5 mM  $CaCl_2$ , 10% (w/v) PEG 2000, 30% (w/v) glycerol). The final concentration of the crystal suspension was determined by measuring Chl concentration of small aliquots of the suspension, dissolved in 80% acetone<sup>49</sup>. The Chl concentration was adjusted between 0.3 and 0.5 mM, corresponding to a protein concentration of 8.5–14 mg ml<sup>-1</sup> (3–5 mg ml<sup>-1</sup>). For solution samples, the purified PS II was resuspended in buffer D (100 mM MES, pH 6.5, 5 mM  $CaCl_2$ , 0.015%  $\beta$ -dodecyl maltoside, 1.3 M sucrose) to a final protein concentration of 80–90 mg ml<sup>-1</sup>.

Thermolysin was obtained from Hampton Research (CA, USA). Microcrystals of thermolysin were obtained as described previously<sup>28</sup> using PEG2000 as a precipitant.

**MIMS measurements.** Sample suspensions of PS II from *T. elongatus* of 8 mg ml<sup>-1</sup> Chl were diluted to 7 mg ml<sup>-1</sup> Chl with H<sub>2</sub><sup>18</sup>O (98%) to give a final enrichment of H<sub>2</sub><sup>18</sup>O of  $\sim 12\%$  and final salt concentrations of 4.4 mM  $CaCl_2$ , 85 mM MES and 1.1 M sucrose. No electron acceptors were added. The <sup>18</sup>O-enriched samples were loaded into a gas-tight Hamilton syringe and pumped through a silica capillary (inner diameter (ID) = 50  $\mu$ m, outside diameter (OD) 160  $\mu$ m) into another gas tight Hamilton syringe that collected the sample. Both syringes were placed on separate syringe pumps. Samples were kept in darkness or very dim green light during all steps, except when illuminated inside the capillary with laser light travelling through one to four optical fibres (400  $\mu$ m core diameter) directly attached to a region of the capillary with the polyimide coating removed. This set-up directly mirrors the in-capillary illumination set up for the CXI experiment (see Fig. 1a and below).

The oscillation pattern of PS II crystals was obtained in the same way, but the experimental details were as follows: the PS II crystal suspension was concentrated to 3.2 mM Chl and was then diluted with H<sub>2</sub><sup>18</sup>O (98%) to 2.5 mM Chl to give a final enrichment of 21.5%. The final concentrations of other additions were 5 mM  $CaCl_2$ , 80 mM MES, 1.2 M sucrose and 11% PEG 2000. The capillary that was used to conduct these experiments had an ID of 100  $\mu$ m and an OD of 360  $\mu$ m.

A Nd:YAG laser (Continuum Inlite II-20, 532 nm, 7 ns pulse width) was used for sample illumination. To obtain a stable output intensity of 7  $\mu$ J per fibre (intensities of individual flashes may vary by  $\pm 5\%$ ), the laser was operated continuously at 20 Hz. The illumination periods were set with the help of a fast shutter (SH05 operated with SC10 Controller; both from Thorlabs), while the flash frequency was controlled via the Q-switch divider (20 or 10 Hz).

The oxygen produced was quantified by injecting the illuminated sample into a membrane-inlet cell containing 600  $\mu$ l water, connected via a Si membrane (Mem 213) and a cooling trap (dry ice/ethanol) to an isotope ratio mass spectrometer (DELTA V, ThermoFinnigan)<sup>50</sup>. The O<sub>2</sub> formed during illumination was detected with excellent S/N ratio as the non-labelled <sup>16</sup>O<sup>16</sup>O, the mixed labelled <sup>16</sup>O<sup>18</sup>O and double-labelled <sup>18</sup>O<sup>18</sup>O species. To obtain a flash pattern, the light-induced yields for O<sub>2</sub> production (detected at  $m/z = 34$ ) obtained with  $(x-1)$  illuminations were subtracted from that with  $x$  illuminations (Fig. 2b). For the first flash, the background <sup>34</sup>O<sub>2</sub> signal of a non-illuminated sample was subtracted. Each measurement was repeated twice (deviation of the points was within 10%).

**Determination of the total O<sub>2</sub> produced per PS II complex.** Determination of the total O<sub>2</sub> produced per PS II complex and flash number requires the absolute calibration of the MIMS signals. This calibration was achieved by the injection of known volumes of air-saturated water into the MIMS cell. This value was used to determine the micromoles of O<sub>2</sub> produced by PS II by the illumination with 3 flashes using 7  $\mu$ J per fibre measured in a silica capillary (ID = 75  $\mu$ m, OD 160  $\mu$ m), with a flow rate of 0.5  $\mu$ l min<sup>-1</sup> and a flash frequency of 4 Hz. To account for diffusion losses of <sup>34</sup>O<sub>2</sub> out of the capillary during the flow of the sample, a loss factor was determined separately by measuring the O<sub>2</sub> content of a PS II sample that was illuminated inside a gas tight syringe by 50 consecutive Xenon lamp flashes (2 Hz frequency, in the presence of acceptors) and either directly injected into the MIMS cell or first flowed through the capillary set-up used for the flash measurements into the collection syringe and then injected into the MIMS cell.

After correcting the O<sub>2</sub> amount obtained in the 3F experiment by the loss factor, this number was then divided by the  $\mu$ mole of PS II reaction centre, which resulted in  $\sim 0.73$  O<sub>2</sub>/RC (Fig. 2a). As three consecutive flashes are required to produce one molecule of oxygen in a dark-adapted PS II reaction centre, a maximum of one O<sub>2</sub>/reaction centre can be expected under these conditions. The above number of 73% oxygen yield directly translates into light saturation. For 100% light saturation, we would expect a value of 73% (0.9<sup>3</sup>), as even under stationary conditions an average 'miss' of 10% occurs due to charge equilibria within PS II.

**Sample injection and illumination at CXI.** Samples were injected into the CXI instrument chamber<sup>32</sup> using an electrospray liquid microjet<sup>28</sup>. Aliquots of 50–150  $\mu$ l of sample were placed in a microcentrifuge tube placed inside the pressurized cell with a Pt-electrode and the end of the injector capillary submerged in the sample. Pressures of 17–20 psi against the CXI chamber pressure (10<sup>-4</sup> Torr) and voltages of around 3,000 V were applied depending on the buffer composition and crystal concentration. The injector capillary was a clear silica capillary with an ID of 75 or 100  $\mu$ m and an OD of 150 or 360  $\mu$ m, respectively. The flow rate was in the range of 0.25–1.0  $\mu$ l min<sup>-1</sup> (for the 75  $\mu$ m ID capillary) and 1.2–3.5  $\mu$ l min<sup>-1</sup> (for the 100  $\mu$ m ID capillary) depending on the sample viscosity and the backing pressure. We measured the flow rate from the mass of sample consumed divided by the run time (data quoted here was obtained this way), and in several cases we also estimated the flow rate from the velocity of crystals flowing in the capillary by using in-capillary visualization of the flow via a microscope-mounted camera. To ensure that the samples were in the dark stable S<sub>1</sub> state before injection, all sample handling and storage was performed in darkness or under dim green light. For visualization of the jet, an infrared laser diode (Coherent Lasiris, 785 nm, 15 mW) was used (see ref. 28 for details). The wavelength was chosen to be outside the absorption spectrum of PS II.

Sample illumination for advancement into higher S-states was conducted using the output of a frequency-doubled Nd:YLF laser at 527 nm (Coherent Evolution) with 150 ns pulse duration. The light from the laser was coupled into multi-mode fibre light guides with a core diameter of 400  $\mu$ m. Three of these light guides (laser 1, 2 and 3 in Fig. 1a) were directly coupled onto the silica capillary of the sample injector and an additional laser (laser 4 in Fig. 1a) illuminated the sample in the jet. The output of all three fibres was equalized to 10  $\mu$ J per pulse (1.4 times the power necessary for saturation of the oxygen production in our offline O<sub>2</sub> flash yield experiments, see above) using wave plates, and the transmission profile and output power of all fibres was measured before the experiment and after each change to the experimental set-up. The centre–centre distance between the fibres was 1.98 mm and typical illumination parameters for a 75- $\mu$ m ID capillary (1  $\mu$ l min<sup>-1</sup> flow rate, 8.9 Hz illumination frequency) result in a delay between the pump laser flashes on the same volume segment of 0.52 s. To generate 2F samples (enriched in the S<sub>3</sub> state), lasers 2 and 3 were used, and the sample reached the X-ray interaction region  $\sim 0.5$  s after the second pump laser flash. To generate 3F samples (enriched in the S<sub>0</sub> state), lasers 1, 2 and 3 were used and the sample reached the X-ray interaction region  $\sim 0.5$  s after the third pump laser flash. Finally, to study transients during the S<sub>3</sub> to S<sub>0</sub> transition, lasers 2 and 3 (for advancing the sample from the S<sub>1</sub> to the S<sub>3</sub> state) in combination with laser 4 (to start the transition from the S<sub>3</sub> to the S<sub>0</sub> state) were used. In this experiment, the time delay of the X-ray probe with respect to the laser 4 was set to 250  $\mu$ s, giving rise to the 3F' data (enriched in the S<sub>3</sub>Y<sub>2</sub><sup>X</sup> state).

**CXI instrument and X-ray parameters.** The CXI instrument of LCLS<sup>32</sup> was used in the 1- $\mu$ m focus setting with the beam being focused to 1.5  $\times$  1.5  $\mu$ m<sup>2</sup> full-width half maximum using Kirkpatrick–Baez mirrors<sup>51</sup>. The pulse length used was  $\sim 45$  fs and the repetition rate was 120 Hz. The energy was varied between 7.1 and 9.7 keV with 3–6  $\times 10^{11}$  photons per pulse. The dose therefore varied between 50 and 300 MGy. XES was measured using 7.1 keV excitation, due to the higher cross section for the Mn transition at this energy. XRD of PS II was measured at both energies but no difference in diffraction quality between the two energies was observed. The majority of the PS II XRD data was obtained with incident energy of 7.1 keV. XRD of thermolysin was collected at 9.7 keV (above the Zn edge).

**X-ray emission spectroscopy setup.** X-ray emission spectra were recorded in a shot-by-shot mode using a custom-built spectrometer in the von Hamos



geometry<sup>25,52</sup>. The crystal analyser array was located 500 mm from the interaction point and with an angle of 81° between the centre of the array and the X-ray beam. A set of 16 Si(440) crystal analysers was used and the Bragg angle range was 85.9° to 83.4°, equivalent to an energy range from 6,473.8 to 6,500 eV (limited by the detector size). The signal was recorded on a 140 k CSPAD<sup>53,54</sup>, located below the interaction region. The set-up was calibrated by recording spectra from solutions of Mn<sup>II</sup>Cl<sub>2</sub> as previously described<sup>25</sup>. The first moment values of the emission spectra were calculated between 6,486 and 6,494 eV as described<sup>35</sup>. Before computing difference spectra, the flash state spectra were smoothed by using either a moving average over a 0.5-eV window or by fitting a cubic polynomial of 3 eV width to each data point.

SR XES spectra were recorded at European Synchrotron Radiation Facility (ESRF) beamline ID 26 using the 440 reflection of five spherically bent ( $r = 1,000$  mm) Si crystal analysers in combination with a silicon drift detector aligned in a Rowland geometry. The overall energy bandwidth of the X-ray emission spectrometer was 0.8 eV. The sample was kept at 15 K in a liquid He cryostat surrounded by He as a heat-exchange gas. The ESRF storage ring was run in 16 bunch mode with ring currents between 60 and 90 mA. The incident beam was monochromatized and tuned to 6.75 keV using the 111 reflection of a pair of cryogenically cooled Si crystals. The beam size was 1.0 (h) × 0.2 (v) mm<sup>2</sup> and the beam position on the sample was changed after 1 s of X-ray illumination.

**Computational facilities.** Over a 5-day period, 114-TB data were collected at LCLS, grouped into five 12-h shifts. Data were processed immediately to assess their completeness and quality. However, as the data size exceeded the processing capacity of the 480-core Linux cluster available at LCLS, arrangements were made to access an additional 1,000 Linux cores at the National Energy Research Scientific Computing Center (NERSC). Transfer of the data from SLAC to NERSC was made over the Energy Sciences Network at a maximum sustained rate of 7.5 Gb s<sup>-1</sup>.

**XRD data processing.** XRD data were recorded using the large CSPAD at LCLS's CXI instrument<sup>32</sup>, and processed using *cctbx.xfel*<sup>55,56</sup>. A dark-current image (pedestal) was subtracted from each image before data reduction. An initial triage step was evaluated, retaining only those images containing 16 or more strong, low-resolution Bragg spots as determined by the *Spotfinder* procedure<sup>56,57</sup>. However, it was found that this step rejected some useful data; thus, it was ultimately omitted from the data processing protocol. Indexing (determination of the unit cell and crystal orientation) was performed with the LABELIT implementation<sup>58</sup> of the Rossmann DPS algorithm<sup>59,60</sup>, and was guided by supplying the known unit cell<sup>61,62</sup>. Where more than one crystal was exposed in the same shot, indexing was attempted on the two most dominant lattices<sup>56</sup>. The number of images or lattices retained after each processing step is detailed in Supplementary Table 1.

Crystal orientations determined by LABELIT were optimized by minimizing the positional difference between the observed Bragg spots and those predicted by the model. Orientational models were further refined so that minimal perturbations were needed to exactly fit the observed Bragg spots to Bragg's law, under the simplifying assumptions of a perfect crystal lattice and a monochromatic beam. Differences between this idealized model and the actual set of observations then allowed us to estimate crystal properties such as mosaicity and the average size of coherently scattering mosaic blocks, leading to a realistic model of Bragg spot positions suitable for signal integration.

Intensities were integrated by summation within a spot mask derived from nearby strong spots atop a planar background<sup>63,64</sup> and corrected for polarization<sup>65</sup>. Intensity variances,  $\sigma^2(I)$ , were derived by counting statistics<sup>66</sup> and a coarse estimate of the detector gain. Error estimates from each diffraction pattern were then inflated by assuming that negative values of  $I/\sigma(I)$  are actually decoy measurements (noise only) with a Gaussian distribution centred at zero and with an s.d. of 1, thus providing a lower bound on modelling errors. A separate resolution cutoff was determined for each image based on a Wilson plot (average intensity versus binned resolution).

Integrated, non-negative intensities from separate images were then scaled to intensities derived from isomorphous reference structures (PDB codes 3bz1 and 3bz2 for PS II<sup>12</sup> and PDB code 2tli for thermolysin<sup>37</sup>), without separately accounting for the partiality fraction of each observation. Images whose intensities correlated poorly ( $\leq 10\%$ ) with those of the reference model were rejected, as were images that deviated from the reference unit cell lengths (10%) or angles (2°), or that did not obey the expected symmetry. Multiple measurements with the same Miller index were merged by averaging and the error was modelled by propagating the  $\sigma(I)$  values in quadrature. The resolution cutoff for the merged data sets was determined from the resolution dependence both of the multiplicity and of  $CC_{1/2}$ , the correlation coefficient of semi-data sets merged from odd- and even-numbered images<sup>36</sup>. The expected contribution of the anomalous signal to  $|F|$  was estimated using the Web server of the Biomolecular Structure Center at University of Washington ([http://skuld.bmsc.washington.edu/scatter/AS\\_index.html](http://skuld.bmsc.washington.edu/scatter/AS_index.html)).

**Phasing and refinement.** As a starting PSII model, we used PDB IDs 3bz1 and 3bz2 (ref. 12, modified to include all atoms in the OEC based on the high-resolution structure 3arc<sup>13</sup>, and re-refined against the 3bz1/3bz2 deposited amplitudes in *phenix.refine*<sup>67</sup>). The structure was then reduced to a single copy of

the PS II complex and the processed data sets were phased by molecular replacement in *Phaser*<sup>68</sup>. Refinement of coordinates and B-factors was performed in *phenix.refine* using tight restraints, including two-fold non-crystallographic symmetry, with the distances between heavy atoms in the OEC restrained to the values determined by EXAFS<sup>20</sup>. Simulated annealing omit maps were generated with the OEC atoms set to zero occupancy, with harmonic restraints<sup>69</sup> applied to the OEC and surrounding atoms; the default parameters of a starting temperature of 5,000 K and 100 K steps were used. Isomorphous difference maps were generated using *phenix.fobs\_minus\_fobs\_map*. Structures and maps were aligned using the *PHENIX* structure comparison tool. All structure figures were created in *PyMOL* 1.2.

The thermolysin structure was solved by molecular replacement using PDB ID 2tli<sup>37</sup> with metals and waters removed, rebuilt using the *PHENIX AutoBuild wizard*<sup>70</sup> and refined in *phenix.refine*. Simulated annealing omit maps were generated as for PS II, with Zn and Ca occupancies set to zero.

## References

- Renger, G. Light induced oxidative water splitting in photosynthesis: energetics, kinetics and mechanism. *J. Photochem. Photobiol. B* **104**, 35–43 (2011).
- Brudvig, G. W. Water oxidation chemistry of photosystem II. *Phil. Trans. R. Soc. B* **363**, 1211–1218 (2008).
- Kok, B., Forbush, B. & McGloin, M. Cooperation of charges in photosynthetic oxygen evolution. I. A linear four step mechanism. *Photochem. Photobiol.* **11**, 457–475 (1970).
- Cox, N. & Messinger, J. Reflections on substrate water and dioxygen formation. *Biochim. Biophys. Acta* **1827**, 1020–1030 (2013).
- 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).
- Yano, J. & Yachandra, V. K. Where water is oxidized to dioxygen: structure of the photosynthetic Mn<sub>4</sub>Ca cluster from X-ray spectroscopy. *Inorg. Chem.* **47**, 1711–1726 (2008).
- Noguchi, T., Suzuki, H., Tsuno, M., Sugiura, M. & Kato, C. Time-resolved infrared detection of the proton and protein dynamics during photosynthetic oxygen evolution. *Biochemistry* **51**, 3205–3214 (2012).
- Haddy, A. EPR spectroscopy of the manganese cluster of photosystem II. *Photosynth. Res.* **92**, 357–368 (2007).
- Cox, N., Pantazis, D. A., Neese, F. & Lubitz, W. Biological water oxidation. *Acc. Chem. Res.* **46**, 1588–1596 (2013).
- Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J. & Iwata, S. Architecture of the photosynthetic oxygen-evolving center. *Science* **303**, 1831–1838 (2004).
- Loll, B., Kern, J., Saenger, W., Zouni, A. & Biesiadka, J. Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II. *Nature* **438**, 1040–1044 (2005).
- 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).
- Umena, Y., Kawakami, K., Shen, J. R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 angstrom. *Nature* **473**, 55–60 (2011).
- Siegbahn, P. E. M. Structures and energetics for O<sub>2</sub> formation in photosystem II. *Acc. Chem. Res.* **42**, 1871–1880 (2009).
- Sproviero, E. M., McEvoy, J. P., Gascon, J. A., Brudvig, G. W. & Batista, V. S. Computational insights into the O<sub>2</sub>-evolving complex of photosystem II. *Photosynth. Res.* **97**, 91–114 (2008).
- Yamaguchi, K. *et al.* Full geometry optimizations of the mixed-valence CaMn<sub>4</sub>O<sub>4</sub>X(H<sub>2</sub>O)<sub>4</sub> (X = OH or O) cluster in OEC of PS II: Degree of symmetry breaking of the labile Mn–X–Mn bond revealed by several hybrid DFT calculations. *Int. J. Quantum Chem.* **113**, 525–541 (2013).
- Haumann, M. *et al.* Photosynthetic O<sub>2</sub> formation tracked by time-resolved X-ray experiments. *Science* **310**, 1019–1021 (2005).
- Yano, J. *et al.* X-ray damage to the Mn<sub>4</sub>Ca complex in single crystals of photosystem II: a case study for metalloprotein crystallography. *Proc. Natl Acad. Sci. USA* **102**, 12047–12052 (2005).
- Grabolle, M., Haumann, M., Müller, C., Liebisch, P. & Dau, H. Rapid loss of structural motifs in the manganese complex of oxygenic photosynthesis by x-ray irradiation at 10–300 K. *J. Biol. Chem.* **281**, 4580–4588 (2006).
- Glöckner, C. *et al.* Structural changes of the oxygen evolving complex in photosystem II during the catalytic cycle. *J. Biol. Chem.* **288**, 22607–22620 (2013).
- Kern, J. *et al.* Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science* **340**, 491–495 (2013).
- Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **406**, 752–757 (2000).
- Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77 (2011).

24. Boutet, S. *et al.* High-resolution protein structure determination by serial femtosecond crystallography. *Science* **337**, 362–364 (2012).
25. Alonso-Mori, R. *et al.* Shot-by-shot energy-dispersive X-ray emission spectroscopy using an X-ray free electron laser. *Proc. Natl Acad. Sci. USA* **109**, 19103–19107 (2012).
26. Glatzel, P. & Bergmann, U. High resolution 1s core hole X-ray spectroscopy in 3d transition metal complexes - electronic and structural information. *Coord. Chem. Rev.* **249**, 65–95 (2005).
27. Kern, J. *et al.* Room temperature femtosecond X-ray diffraction of photosystem II microcrystals. *Proc. Natl Acad. Sci. USA* **109**, 9721–9726 (2012).
28. Sierra, R. G. *et al.* Nanoflow electrospinning serial femtosecond crystallography. *Acta Cryst. D* **68**, 1584–1587 (2012).
29. de Wijn, R. & van Gorkom, H. J. Kinetics of electron transfer from Q(a) to Q(b) in photosystem II. *Biochemistry* **40**, 11912–11922 (2001).
30. Krivanek, R., Kern, J., Zouni, A., Dau, H. & Haumann, M. Spare quinones in the Q(B) cavity of crystallized photosystem II of *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta* **1767**, 520–527 (2007).
31. Shevela, D. & Messinger, J. Probing the turnover efficiency of photosystem II membrane fragments with different electron acceptors. *Biochim. Biophys. Acta* **1817**, 1208–1212 (2012).
32. Boutet, S. & J Williams, G. The Coherent X-ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). *New J. Phys.* **12**, 035024 (2010).
33. Iuzzolino, L., Dittmer, J., Dörner, W., Meyer-Klaucke, W. & Dau, H. X-ray absorption spectroscopy on layered photosystem II membrane particles suggests manganese-centered oxidation of the oxygen-evolving complex for the S<sub>0</sub>-S<sub>1</sub>, S<sub>1</sub>-S<sub>2</sub>, and S<sub>2</sub>-S<sub>3</sub> transitions of the water oxidation cycle. *Biochemistry* **37**, 17112–17119 (1998).
34. Kulik, L. V., Epel, B., Lubitz, W. & Messinger, J. Electronic structure of the Mn<sub>4</sub>O<sub>x</sub>Ca cluster in the S<sub>0</sub> and S<sub>2</sub> states of the oxygen-evolving complex of photosystem II based on pulse Mn<sup>55</sup>-ENDOR and EPR spectroscopy. *J. Am. Chem. Soc.* **129**, 13421–13435 (2007).
35. Messinger, J. *et al.* Absence of Mn-centered oxidation in the S<sub>2</sub> -> S<sub>3</sub> transition: implications for the mechanism of photosynthetic water oxidation. *J. Am. Chem. Soc.* **123**, 7804–7820 (2001).
36. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
37. English, A. C., Done, S. H., Caves, L. S. D., Groom, C. R. & Hubbard, R. E. Locating interaction sites on proteins: the crystal structure of thermolysin soaked in 2% to 100% isopropanol. *Proteins* **37**, 628–640 (1999).
38. Barends, T. R. *et al.* Anomalous signal from S atoms in protein crystallographic data from an X-ray free-electron laser. *Acta Cryst. D* **69**, 838–842 (2013).
39. Barends, T. R. *et al.* De novo protein crystal structure determination from X-ray free-electron laser data. *Nature* **505**, 244–247 (2013).
40. Ferrer, J. L., Larive, N. A., Bowler, M. W. & Nurizzo, D. Recent progress in robot-based systems for crystallography and their contribution to drug discovery. *Expert Opin. Drug Discov.* **8**, 835–847 (2013).
41. Zouni, A. *et al.* Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 angstrom resolution. *Nature* **409**, 739–743 (2001).
42. Biesiadka, J., Loll, B., Kern, J., Irrgang, K. D. & Zouni, A. Crystal structure of cyanobacterial photosystem II at 3.2 angstrom resolution: a closer look at the Mn-cluster. *Phys. Chem. Chem. Phys.* **6**, 4733–4736 (2004).
43. Kamiya, N. & Shen, J. R. Crystal structure of oxygen-evolving photosystem II from *Thermosynechococcus vulcanus* at 3.7-Å resolution. *Proc. Natl Acad. Sci. USA* **100**, 98–103 (2003).
44. Kern, J. *et al.* Cyanobacterial photosystem II at 3.2 Å resolution - the quinone binding pockets. *Photosynth. Res.* **84**, 153–159 (2005).
45. Rappaport, F., Blanchardesce, M. & Lavergne, J. Kinetics of electron-transfer and electrochromic change during the redox transitions of the photosynthetic oxygen-evolving complex. *Biochim. Biophys. Acta* **1184**, 178–192 (1994).
46. Rappaport, F., Ishida, N., Sugiura, M. & Boussac, A. Ca<sup>2+</sup> determines the entropy changes associated with the formation of transition states during water oxidation by photosystem II. *Energ. Environ. Sci.* **4**, 2520–2524 (2011).
47. Razeghifard, M. R. & Pace, R. J. EPR kinetic studies of oxygen release in thylakoids in PSII membranes: A kinetic intermediate in the S<sub>3</sub> to S<sub>0</sub> transition. *Biochemistry* **38**, 1252–1257 (1999).
48. Kern, J. *et al.* Purification, characterisation and crystallisation of photosystem II from *Thermosynechococcus elongatus* cultivated in a new type of photobioreactor. *Biochim. Biophys. Acta* **1706**, 147–157 (2005).
49. Porra, R. J., Thompson, W. A. & Kriedemann, P. E. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta* **975**, 384–394 (1989).
50. Beckmann, K., Messinger, J., Badger, M. R., Wydrzynski, T. & Hillier, W. On-line mass spectrometry: membrane inlet sampling. *Photosynth. Res.* **102**, 511–522 (2009).
51. Siewert, F. *et al.* Ultra-precise characterization of LCLS hard X-ray focusing mirrors by high resolution slope measuring deflectometry. *Opt. Express* **20**, 4525–4536 (2012).
52. Alonso-Mori, R. *et al.* A multcrystal wavelength dispersive X-ray spectrometer. *Rev. Sci. Instrum.* **83**, 073114 (2012).
53. Hart, P. *et al.* The CSPAD megapixel x-ray camera at LCLS. *Proc. SPIE* **8504**, 85040C–85041C (2012).
54. Herrmann, S. *et al.* Cspad-140k: a versatile detector for LCLS experiments. *Nucl. Instrum. Meth. A* **718**, 550–553 (2013).
55. Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. New Python-based methods for data processing. *Acta Cryst. D* **69**, 1274–1282 (2013).
56. Hattne, J. *et al.* Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nat. Methods* **11**, 545–548 (2014).
57. Zhang, Z., Sauter, N. K., van den Bedem, H., Snell, G. & Deacon, A. M. Automated diffraction image analysis and spot searching for high-throughput crystal screening. *J. Appl. Crystallogr.* **39**, 112–119 (2006).
58. Sauter, N. K., Grosse-Kunstleve, R. W. & Adams, P. D. Robust indexing for automatic data collection. *J. Appl. Crystallogr.* **37**, 399–409 (2004).
59. Steller, I., Bolotovskiy, R. & Rossmann, M. G. An algorithm for automatic indexing of oscillation images using Fourier analysis. *J. Appl. Crystallogr.* **30**, 1036–1040 (1997).
60. Rossmann, M. G. & van Beek, C. G. Data processing. *Acta Cryst. D* **55**, 1631–1640 (1999).
61. Paithankar, K. S. *et al.* Simultaneous X-ray diffraction from multiple single crystals of macromolecules. *Acta Cryst. D* **67**, 608–618 (2011).
62. White, T. A. *et al.* Crystallographic data processing for free-electron laser sources. *Acta Cryst. D* **69**, 1231–1240 (2013).
63. Otwinowski, Z. & Minor, W. *Processing of X-Ray Diffraction Data Collected in Oscillation Mode* Vol. 276 Macromolecular Crystallography, Part A (Academic Press, 1997).
64. Leslie, A. G. Integration of macromolecular diffraction data. *Acta Cryst. D* **55**, 1696–1702 (1999).
65. Kahn, R. *et al.* Macromolecular crystallography with synchrotron radiation - photographic data-collection and polarization correction. *J. Appl. Crystallogr.* **15**, 330–337 (1982).
66. Leslie, A. G. The integration of macromolecular diffraction data. *Acta Cryst. D* **62**, 48–57 (2006).
67. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst. D* **68**, 352–367 (2012).
68. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
69. Hodel, A., Kim, S. H. & Brunger, A. T. Model bias in macromolecular crystal structures. *Acta Cryst. A* **48**, 851–858 (1992).
70. Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Cryst. D* **64**, 61–69 (2008).

## Acknowledgements

This work was supported by NIH Grant GM055302 (V.K.Y.) for PS II biochemistry, structure and mechanism; the Director, Office of Science, Office of Basic Energy Sciences (OBES), Division of Chemical Sciences, Geosciences and Biosciences (CSGB) of the Department of Energy (DOE) under Contract DE-AC02-05CH11231 (J.Y. and V.K.Y.) for X-ray methodology and instrumentation; by NIH grant P41GM103393 for part of the XES instrumentation and support of U.B.; an LBNL Laboratory Directed Research and Development award to N.K.S.; and NIH grants GM095887 and GM102520 (N.K.S.) for data processing methods. The DFG-Cluster of Excellence ‘UniCat’ coordinated by the Technische Universität Berlin and Sfb1078, TP A5 (A.Z.); the Alexander von Humboldt Foundation (J.K.); the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement no 283745 (CRISP) (P.G.); the Solar Fuels Strong Research Environment (Umeå University), the Artificial Leaf Project (K&A Wallenberg Foundation), Energimyndigheten (J.M.) and VR (M.M.S., J.M.), and the Human Frontiers Science Project (J.Y., U.B. and A.Z.) are acknowledged for supporting this project. The injector work was supported by DOE Office of Basic Energy Sciences, Chemical Sciences Division, under Contract DE-AC02-76SF00515 (H.L., C.A.S. and M.J.B.), LCLS (R.G.S.), the Human Frontiers Science Project Award RPG005/2011 (H.L.) and the SLAC Laboratory Directed Research and Development Program (C.A.S. and M.J.B.). We thank Professor Ken Sauer for continuing scientific discussions. We thank Tom Terwilliger, Randy Read, Nigel Moriarty, Ralf Grosse-Kunstleve, Pavel Afonine and Jeffrey Headd for helpful discussions and technical assistance regarding XRD data processing; Matthew Latimer for support with the XES set-up; Don Schaefer, Alan Miahnahri and William White for support with development of the laser illumination and injector set-up; Christopher Kenney, Ryan Herbst, Jack Pines, Philip Hart, John Morse, Gunther Haller and Sven Herrmann for support with the CSPAD detectors; Amedeo Perazzo and Igor Gaponenko (SLAC) for computing support; Gregory Bell (Energy Sciences Network (ESnet)) for arranging network access for data transfer; and Shane Canon (NERSC) for arranging computing access. We thank the staff at LCLS/SLAC and the staff at SSRL, ALS, APS and ESRF for support of synchrotron experiments. Portions of this research

were carried out at the Linac Coherent Light Source (LCLS) at the SLAC National Accelerator Laboratory. LCLS is an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Stanford University. Data processing was performed in part at the National Energy Research Scientific Computing Center, supported by the DOE Office of Science, contract number DE-AC02-05CH11231. Testing of crystals and various parts of the set-up were carried out at synchrotron facilities that were provided by the Advanced Light Source (ALS), BL 5.0.2, in Berkeley, and Stanford Synchrotron Radiation Lightsource (SSRL), BL 6-2, in Stanford, funded by DOE OBES. The SSRL Biomedical Technology programme is supported by NIH, the National Center for Research Resources and the DOE Office of Biological and Environmental Research. SR PS II XES data were recorded at the European Synchrotron Radiation Facility (ESRF), Grenoble, France.

### Author contributions

U.B., V.K.Y. and J.Y. conceived the experiment; U.B., J.Y., V.K.Y., J.K., R.A.-M., J.M., A.Z., N.K.S., G.J.W., S.B., A.R.F., D.M. and M.J.B. designed the experiment; R.T., J. Hellmich, G.H., R.C., D.D., M.I., C.G., J.K., A.L., B.L.-K., S.G. and A.Z. prepared samples; S.B., J.E.K., M.M., M.M.S. and G.J.W. operated the CXI instrument; M.J.B., H.L., R.G.S., J.K., J.M., B.L.-K., S.G., R.T. and C.G., J.Hellmich and G.J.W. developed, tested and ran sample delivery system; S.K. and J.M. conceived, set up and performed O<sub>2</sub> evolution measurements; R.C., R.A.-M., E.G., J.U. and P.G. collected SR XES of PS II; R.A.-M., J.K., R.T., B.L.-K., S.G., T.-C.W., D.S. and J.Y. developed and tested XES set-up; R.A.-M., A.B., U.B., M.J.B., S.B., R.C., R.J.G., P.G., C.G., S.G., G.H., J.Hattne., J.Hellmich,

J.K., J.E.K., A.L., H.L., B.L.-K., D.M., M.M., J.M., N.K.S., M.M.S., R.G.S., C.A.S., D.S., R.T., T.-C.W., G.J.W., V.K.Y., J.Y. and A.Z. performed the LCLS experiment; J. Hattne, N.E., R.J.G., A.B., R.A.-M., J.K., C.A.S., P.H.Z., M.M., P.D.A. and N.K.S. developed new software and/or processed and analysed data; D.E.S. arranged computer access; J.K., J.Y., J.M., N.K.S., U.B. and V.K.Y. wrote the manuscript with input from all authors.

### Additional information

**Accession codes:** The X-ray crystallographic coordinates and structure factors for structures reported in this study have been deposited at the Protein Data Bank (PDB), under deposition numbers 4TNL (thermolysin), 4TNH (PS II dark state), 4TNJ (PS II 2F), 4TNI (PS II 3F) and 4TNK (PS II 3F'). These data can be obtained free of charge from [www.pdb.org](http://www.pdb.org).

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

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

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Kern, J. *et al.* Taking snapshots of photosynthetic water oxidation using femtosecond X-ray diffraction and spectroscopy. *Nat. Commun.* 5:4371 doi: 10.1038/ncomms5371 (2014).