PNAS
www.pnas.org

Supplementary Information for

Untangling the Sequence of Events During the $S_2 \rightarrow S_3$ Transition in Photosystem II and Implications for the Water Oxidation Mechanism

Mohamed Ibrahim[‡], Thomas Fransson[‡,], Ruchira Chatterjee[‡], Mun Hon Cheah[‡], Rana Hussein, Louise Lassalle, Kyle D. Sutherlin, Iris D. Young, Franklin D. Fuller, Sheraz Gul, In-Sik Kim, Philipp S. Simon, Casper de Lichtenberg, Petko Chernev, Isabel Bogacz, Cindy C. Pham, Allen M. Orville, Nicholas Saichek, Trent Northen, Alexander Batyuk, Sergio Carbajo, Roberto Alonso-Mori, Kensuke Tono, Shigeki Owada, Asmit Bhowmick, Robert Bolotovsky, Derek Mendez, Nigel W. Moriarty, James M. Holton, Holger Dobbek, Aaron S. Brewster, Paul D. Adams, Nicholas K. Sauter, Uwe Bergmann, Athina Zouni[*], Johannes Messinger[*], Jan Kern, Vittal K. Yachandra[*], Junko Yano[*]

Emails: Athina.Zouni@hu-berlin.de, johannes.messinger@kemi.uu.se, VKYachandra@lbl.gov, JYano@lbl.gov

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.2000529117

Supplementary Information Text

Methods

Sample Preparation: PS II dimers were extracted and purified from *T. elongatus* using a modified protocol reported previously (1). PS II crystals ranging in size from 20-60 μm were prepared using a seeding protocol (2). The crystals were dehydrated by treatment with PEG 5000 (1). The 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, at ~0.5-0.8 mM chlorophyll concentration. PS II crystal suspensions for XRD were loaded into the sample delivery syringe (Hamilton gastight syringe, 1 ml) followed by dark-adaption for 1 hour prior to data collection. Solution PS II samples for XES measurements were in a buffer comprising, 0.1 M MES, pH 6.5, 5 mM calcium chloride and 0.015% (w/v) β-dodecyl maltoside (βDM) at 4.0-5.0 mM chlorophyll concentration (40-50 mg/ml protein). PS II solution samples were loaded into the sample delivery syringe (Hamilton gastight syringe, 1 or 2.5 mL) followed by a pre-illumination treatment using green LED diodes (525nm, Thorlabs, USA). The intensity of the light used was 2 μmol $m^{-2}s^{-1}$, the samples were exposed for 10 s while rotating the syringe, followed by dark adaption for 40 mins prior to data collection. We note that the PS II core complexes in our sample preparation contain sufficient number of natural quinones to drive the catalytic reaction through the cycle (3-5).

Characterization: *O_2 evolution rate:* Measurement of steady state O_2 evolution rate under continuous illumination was performed using a Clarke-type electrode. PS II crystals dissolved in a buffer (0.1 M MES, pH 6.5, 0.1 M ammonium chloride, 0.4 mM PPBQ) show O_2 evolution rates of 2520 ± 90 μmol O_2/(mg(Chl) × h), which is 90% of the activity before crystallization. PS II solution samples for XES show O_2 evolution rates of 3000 ± 200 μmol O_2/(mg(Chl) × h) when measured in buffer containing 20 mM MES, pH 6.5, 20 mM calcium chloride, 10 mM magnesium chloride and 0.4 mM PPBQ, indicating approximately 84% of activity is retained after crystallization. The difference in rate is most likely due to different buffer conditions employed, which leads to different turnover kinetics, for measurements of crystals suspensions versus solution samples.
EPR measurements: All sample batches used at the LCLS or SACLA were monitored for Mn(II) content using low-temperature X-band EPR spectra. The Mn(II) content was measured at 20 K, using a Varian E109 EPR spectrometer equipped with a Model 102 Microwave bridge, and the sensitivity of the measurement allowed us 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 is in agreement with that determined by *in situ* XES measurements (6).
MIMS measurements: The S-state advancement of crystals was evaluated by MIMS (membrane inlet mass spectroscopy). A crystal suspension with approximately 10% ^{18}O-labelled water was placed in a thin-layer MIMS setup (7, 8). The sample was given one laser pre-flash and dark adapted at least 40 minutes at room temperature. The sample was then subjected to laser flashes at 5 Hz frequency and O_2 yield was monitored as a peak at m/z 34. Due to slow mass transport of evolved O_2 from the MIMS cell to the ion source of the mass spectrometer, O_2 yield from each individual laser flash cannot be resolved. Instead, the total O_2 yield for 2-flash (2F) is detected and the procedure is repeated for 3F and 4F using a new volume of crystal suspension each time. The O_2 yield pattern as a function of flash number was calculated by subtracting the O_2 yield of a flash number from the yield obtained for the preceding flash number. The miss parameter was determined to be 13% for crystal suspensions and 11% for solution samples.
Determination of the S-state population: We evaluated the S-state advancement of PS II crystal suspensions by two methods, *in situ* and *ex situ*. X-ray emission spectroscopy that monitors the oxidation state of Mn was collected simultaneously with the XRD data as described previously (6, 9, 10). Flash-induced oxygen evolution measurements (MIMS) for O_2 detection (7, 11, 12) was carried out prior to the XFEL experiment. We have estimated the S-state population from MIMS measurements that are shown in Table S2. For the purpose of refinement of the structural models, details regarding the refinement of the mixed models in the 1F and 2F data are given below in the section, “Model building and map calculation”.

Sample injection and illumination: Acoustic droplet ejection (ADE) (13) was used in combination with the Drop-on-Tape (DOT) sample delivery method (10). For capturing the stable intermediates S_2, S_3, and S_0, each droplet of the crystal suspension was illuminated by 120 ns laser pulses at 527 nm using a Nd:YLF laser (Evolution, Coherent) for LCLS or by 8 ns laser pulses at 532 nm using a Nd;YAG laser (Minilite, Continuum) for SACLA via fiber-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, similar to what was used previously (12). We implemented a feedback control system of the belt speed and deposition delay, and the flashing delay and droplet phase were adjusted accordingly (10). To achieve time delays shorter than 200 ms between illumination and X-ray probe a second "free space" laser was utilized. This was either an Opolette 355 LD laser (Opotek, 530 nm wavelength, 7 ns pulse width) at the MFX instrument or an Evolution Nd:YLF laser (Coherent, 527 nm wavelength, 120 ns pulse width) at the XPP instrument. This "free space" laser was triggered to be synchronized with the X-ray pulse with an adjustable delay that was set between 50 and 730 μs for this study. The laser was guided with optics to the X-ray interaction spot and its position was fine tuned for each delay time to ensure that the laser spot position coincides with the position of the sample droplet at the selected delay timing. At the XFELs, a light intensity of 120±10 mJ/cm^2 was applied as O_2 evolution was found to be saturated at 70 mJ/cm^2 for the dimensions and concentrations of samples, both crystals and solutions, used in our experiments (12). A light intensity of 120 mJ/cm^2 corresponds to about 140 photons/PS II monomer and given a minimum pulse length of 7 ns and 35 Chl per PS II monomer this averages to 0.6 photons/(Chl and ns). If a PS II center is undergoing charge separation additional photons absorbed by the internal antenna Chl are rapidly dissipated in form of fluorescence with an average fluorescence lifetime of ~0.5-1 ns, hence preventing any overexcitation of the reaction center.

X-ray data collection: The crystallography data were collected at the MFX instrument of LCLS at SLAC National Accelerator Laboratory, Stanford (14, 15), and at the BL2-EH3 of SACLA, Japan (16, 17), solution samples were collected at the XPP instrument of LCLS (18). XRD and XES of PS II crystals was measured using X-ray pulses of ~40 fs length at 9.5 keV with pulse energies of 2-4 mJ and with an X-ray spot size at the sample of ~3 μm in diameter at LCLS and ~7 fs at 9.5 keV with pulse energies of 300-450 μJ and a spot size of 2x2.5 μm^2 at SACLA. XRD data were collected on a Rayonix MX340-XFEL detector operating in the 3-by-3 binning mode at a frame rate of 20 Hz at the LCLS. This mode provided the optimal tradeoff of resolving power between adjacent Bragg reflections and quantity of images collected. At SACLA an octal MPCCD detector operating at 30 Hz was used. XES data collection from solutions used pulses of <40 fs length at 9.65 keV with pulse energies of 2-2.5 mJ and a spot size of ~3 μm in diameter and a repetition rate of 60 Hz at LCLS.

Mn Kβ X-ray emission spectra (XES) from PS II solutions and crystals were collected using a van Hamos geometry. An array of 16 analyzer crystals with a focal length of 500 mm was placed perpendicular to the beam as described previously (6, 9, 19, 20) and the signal was collected either on an ePIX 100 detector (LCLS) or a single module MPCCD detector (SACLA) mounted underneath the interaction point between the sample and the X-rays. For solution samples the data collection speed was increased to 60 Hz (due to the faster readout of the ePIX detector compared to the Rayonix XRD detector), resulting in a shorter spacing between individual sample droplets on the Kapton tape and hence a small amount of cross illumination (3%) between adjacent drops is observed that is absent in the crystal measurements.

X-ray diffraction data processing: The *cctbx.xfel* graphical user interface was used 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 (21-28). 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 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, ~85% of crystal hits could be indexed.

One of the significant challenges in serial crystallography at XFELs is to accurately model the detector position, including detector distance and beam center, since both of these tend to drift during an experiment. Uncertainty in these parameters will introduce error in spot prediction, which will add noise and reduce signal in the merged dataset. It will also contribute to introducing an artificial non-isomorphism of the unit cell, and lead to images being rejected due to their unit cells being outside an acceptable range for isomorphous data, distorting the quality of the final electron density map (26).
To account for these issues, we carefully refined the detector position prior to integration, according to the procedures in (26) (see also (29)). First, the powder diffraction pattern of a silver(I) behenate sample (Alfa Aesar) was used to obtain an initial estimate of detector distance. This distance was used to index a subset of data using *dials.stills_process* with a target unit cell of a=117.0 Å, b=221.0 Å, c=309.0 Å, $\alpha=\beta=\gamma=90°$ and the space group $P2_12_12_1$. These results were used to further refine the detector position. The higher-precision detector positions were used in subsequent indexing and integration trials for the full dataset, resulting in a maximum of four distinct lattices indexed on a single shot. Next, these indexing and integration results were further improved by performing ensemble refinement of the crystal and detector parameters using *cctbx.xfel_stripe_experiment*. This program batches the data according to when it was collected and refines the detector model separately for each sample batch. These time-sensitive corrections improved the unit cell distribution and final isomorphous difference maps.
After ensemble refinement of the detector and crystal parameters, 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 (10). After filtering out lattices that belonged to a different crystal isoform, a total of 229,810 integrated lattices was obtained with an average unit cell of a=117.0 Å, b=221.7 Å, c=307.6 Å, $\alpha=\beta=\gamma=90°$ and the space group $P2_12_12_1$. These integrated images were merged using *cxi.merge* as described previously (12), with a couple of modifications. As the average unit cell obtained in this experiment was slightly different from what was obtained previously (12), a fully merged dataset at 1.90 Å, containing images extending to at least 2.5 Å from all illumination states, was first obtained using a previously obtained reference model with no restrictions on the unit cell parameters (data not shown). These combined structure factors were used to refine an atomic coordinate model as described below, and this was used as the reference model for merging the separate illumination states. The unit cell outlier rejection mechanism in *cxi.merge* was used to remove images with a unit cell that differed by more than 1% from the reference model, so a pre-filtering step was not necessary. Using a 1% unit cell tolerance was previously found to be essential to reduce noise in isomorphous difference maps, and we found that using such a strict filter did not significantly reduce the final resolutions (12). For the final datasets, only images with reflections extending to at least 3 Å were included, as we have done previously for PSII datasets to improve statistics by removing contamination due to low-quality images (12).

After applying the unit cell and resolution filters, final merged datasets were acquired for the 0F, 1F, 2F(50 μs), 2F(150 μs), 2F(250 μs), 2F(400 μs), and 2F states to resolutions between 2.27 and 2.01 Å, containing between 4464 and 11734 images (Table S1).

Data quality evaluation: The resolution limits reported for each data set in Table S1 were determined by a multistep process. Fundamentally for each separate image and lattice, we assessed the outer resolution bin where Bragg spot intensities were detectable. After using standard 2D summation integration (30) to obtain the Bragg spot intensities I and standard deviations $\sigma(I)$ for one lattice, we cut the data off at the first bin where the average $I/\sigma(I)$ ratio fell below 0.1. That is, even though the lattice periodicity allows us to predict spot positions to the corner of the detector, the detected signal does not warrant including all the data for every image.

Such per-image resolution cutoffs are illustrated in Fig. S2, where panels b,d show three images where differing cutoffs are applied. After postrefinement and scaling (21), duplicate intensity measurements were averaged using weights based on the Ha14 algorithm for intensity variance calculation (31). Agreement of semi-datasets is then quantified using $CC_{1/2}$ (28), and the outer limit defined as that resolution bin where $CC_{1/2}$ either reaches zero, or no longer decreases monotonically (indicating there is no useful information beyond this point). The cutoff is also confirmed using paired refinements as suggested by Karplus and Diederichs (32).

We note that photosystem II papers from other groups (e.g. (33, 34)) have assumed, without justification, that high data quality arises from high multiplicity of coverage. Such logic is not valid, because the multiplicity level is entirely determined by two choices related to spot prediction: 1) the resolution cutoff discussed above, and 2) the allowable offset from the Ewald sphere, which is a presumption of how far reciprocal lattice points can be from the ideal Bragg's law diffraction condition (compare Fig. S2, panels a and c), also see Fig. S3. High multiplicity is therefore not sufficient for data quality, for one must also insure that the predicted spots actually have Bragg-spot photons recorded on their pixels. Conceptually, if the claim is to be made that data quantity equates with data quality, then the $I/\sigma(I)$ ratio should be inspected for unmerged spot predictions on each image, both as a function of resolution and of Ewald offset, as suggested by Fig. S2.

The cautions we raise here are independent of whether data processing is performed with *cctbx.xfel*, as done for Table S1, or with *CrystFEL* (35, 36). In particular neither *cctbx.xfel* nor *CrystFEL* confer automatic assurance that multiplicity is a reliable indication of data quality. To explore further, we processed one of our datasets (2F(250 μs)) with different spot prediction settings using both programs (data not shown). Specifically, we found that both programs permit the choice of whether to use standard cutoffs for resolution limit and Ewald offset across all images, or to customize resolution cutoffs according to the signal strength of each image separately (indexamagig's *integration-rescut* option in *CrystFEL* or the *cctbx.xfel.merge*'s *select.algorithm=significance_filter* option). These spot prediction choices strongly affect multiplicity of coverage, but importantly, the multiplicity is similar when equivalent choices are made for each program. However, the software settings do not strongly affect the science results. Molecular models refined to similar R-factors, and omit maps produced approximately the same peak height (Fig. S4) although with different levels of noise visible in the maps. Method-dependent differences between the results warrant further examination, but detailed analysis is not performed here.

XES data processing and kinetics analysis: Single crystal XES data collected simultaneously with the diffraction images were used to sort out and exclude any sample batches that indicate the presence of Mn(II) released during the on-site crystallization, as well as to confirm the advancement of S-states by fiber-coupled lasers and a free space laser. Details on the experimental setup and data analysis tools can be found in Ref. (6). For the analysis of $S_2 \rightarrow S_3$ transition kinetics, PS II solution data were collected, as this provides more efficient data collection than using single crystal samples (see "XES data of crystals and solution samples"). The compatibility of solution versus crystal X-ray emission spectra, as well as of spectra collected using different pulse durations are discussed below (Figs. S7, S8).

For the analysis of the kinetics of the $S_2 \rightarrow S_3$ transition, we measured first moments for five separate time points (50, 150, 250, 400, and 730 μs) to test the performance of different transition models and time constants. Here, the population of S_1, S_2, S_3, and S_0 were calculated at times following the second flash, using population estimates and miss and cross-illumination parameters discussed above. Due to the cross-illuminations and S_0 population in the dark state also $S_0 \rightarrow S_1$ and $S_3 \rightarrow S_0$ transitions had to be included in the analysis. These, as well as $S_1 \rightarrow S_2$ transition models and time constants, were adopted from the time-dependent X-ray absorption spectroscopy (TDXAS) study of Dau and coworkers (37). Here, the $S_0 \rightarrow S_1$ and $S_1 \rightarrow S_2$ transitions were described as one-step (exponential) electron transfer processes, featuring time constants of 52 and 89 μs, respectively. The $S_3 \rightarrow S_0$ process was reported as being composed of a lag period of 153 μs, followed by a transition to S_0 with a time constant of 1538 μs. The transition to S_0 is

modeled here by the sequential process of $S_3 \rightarrow [S_{2\,like}] \rightarrow S_0$, such that the OEC proceeds through a reduced intermediate state with the same first moment as S_2 before being transferred to full S_0.

For the $S_2 \rightarrow S_3$ transition, Fig. 5 includes the first moment progression obtained by using the transition models and time constants from the TDXAS and TDIR studies of Dau and Noguchi, respectively (37, 38). From TDXAS the process is described as one with a single lag phase, which is again here modeled as $S_2 \rightarrow [S_2] \rightarrow S_3$, using time constants of 26 and 317 µs. The TDIR instead reported two phases before the final transition to S_3, but due to the limited time resolution here we do not include the first fast lag (13 µs), and thus model it as $S_2 \rightarrow [S_2] \rightarrow S_3$, using reported time constants of 104 and 352 µs. Fig. S9 shows fits to our XES data using either one single process with τ_1=350 µs (Fig. S9a) or two time constants τ_1 = 50 µs and τ_2 =300 µs (Fig. S9b). Populations of the individual S-states during the $S_2 \rightarrow S_3$ transition are shown using the same kinetic models and the starting populations for the 1F state as well as miss and double hit parameters either for solution (Fig. S9c) or crystal samples (Fig. S9d).

Model building and map calculation: Initial structure refinement against the 'combined' dataset at 1.90 Å was carried out starting from a previously acquired high-resolution PS II structure in the same unit cell (PDB ID: 5TIS)(3) using *phenix.refine* (39, 40). B-factors were reset to a value of 30 and waters were removed. 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 σ 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-heme iron, and the OEC. Following real space refinement in *Coot* (41) 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' dataset model was obtained with R_{work}/R_{free} of 19.17%/22.42%.

The above model, with reset B-factors to 30 and removed waters, was subsequently refined against the illuminated datasets with the lattermost refinement settings and different OEC bonding restraints. OEC bonding restraints for the 0F dataset prevented large deviations from the high-resolution dark state OEC structure reported by Suga *et al.* (PDB ID: 4UB6 (42)). Bonding restraints for the other datasets loosely restrained the models to metal-metal distances matching spectroscopic data and metal-oxygen distances matching the most likely proposed models (43-47). 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 of cycles of refinement in this manner, individual illuminated states at various resolutions were obtained ranging in R_{work}/R_{free} from 18.17%/23.13% to 19.14%/23.59% (Table S1).

As before (12), after the initial refinement cycles, a positive peak was visible in the mF_{obs}-DF_{calc} density close to Mn1 for the 2F400µs and 2F illumination states, which we attribute to an additional oxygen present in the S_3 state, O_X. The O_X density is also visible in the 2F250µs timepoint, and due to the improved resolution of the current dataset relative to the one we reported previously (12), we now also see the O_X density in the 2F150µs data. This atom was included in the model and in the OEC CIF restraints in the final refinement for these four states. To best approximate the contributions of dimers that did not advance to the next S-state due to illumination misses, for the 2F(150µs), 2F(250µs), 2F(400µs), and 2F datasets, the 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 A54-65, A159-191, A327-344, C327-329, C353-359, D351-352. Population of the S_2 and S_3 states in the data was estimated based on oxygen

evolution and XES measurements (Table S2) and rounded to the nearest 5%, yielding a S_2 state population of 90% in the 1F crystal data and a S_3 state population of 75% in the 2F crystal data. For the time points of the S_2-S_3 transition the population of the S_3 state was estimated based on the observed O_X omit map density and the measured XES kinetics and subsequently rounded to the nearest 5%, yielding values of 30%, 50%, and 60%, for the 2F(150us), 2F(250us), and 2F(400us) datasets, respectively. Accordingly, for these datasets the main conformer across this entire region was set at 0.3 (2F(150us)), 0.5 (2F(250us)), 0.6 (2F(400us)), and 0.75 (2F(200ms)) occupancy and the minor conformer was set to give a total occupancy of 1. The major conformer was allowed to refine as usual, while the minor conformer was fixed during refinement and set to match the S_2 state obtained from the refined 1F structure. This was achieved by least-squares fitting the refined S_2 model onto the new model at the split region in *PyMol* (48) 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 modeling 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 dataset.

Estimated positional precision: The maximum-likelihood coordinate error calculated during refinement is a general-purpose metric for positional error; however, this is an overall metric that gives no information on specific bonds and angles. To estimate the positional precision of the OEC atoms, we used *END/RAPID* to perturb the structure factors, in an approach similar to one we previously employed (49). The structure factors were adjusted by $\pm|F_{obs} - F_{model}|$ in 100 trials using the *END/RAPID* command line tools (https://bl831.als.lbl.gov/END/RAPID/end.rapid/Documentation/end.rapid.Manual.htm), adding noise proportional to the error in the model to generate 100 perturbed datasets for each illumination state. We then re-refined the models against each new dataset, kicking the initial atom positions in the model using the *sites.shake* tool in Phenix before refinement. In this case, both conformers were allowed to refine to accommodate the perturbed data. The mean and standard deviation of selected bond distances were calculated across the re-refined models (49). Depending on resolution and the fraction of S_3 present, metal-metal distances at the OEC had standard deviations between 0.07 and 0.27 Å across these trials, while distances between OEC metals and coordinating ligands were found to have standard deviations between 0.09 and 0.25 Å.

Estimating the uncertainty of the omit densities: Changes in the electron density at the O_X position were obtained from O_X 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 datasets was also used to estimate the uncertainty of the normalized omit density. The omit densities for each dataset were divided by the average omit densities of the chloride ions in that dataset to equate the densities between different illumination states.

Isomorphous difference maps: Isomorphous difference maps for various illumination states were generated using *phenix.fobs_minus_fobs_map*. As before (12), to minimize artifacts in the difference maps due to slight differences between the unit cells of the illumination states, images that differed from the reference unit cell by more than 1% were excluded from merging, as described above.

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.

Data availability: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 6W1O for the 0F, 6W1P for the 1F, 6W1Q for the

2F(50 μs), 6W1R for the 2F(150 μs), 6W1T for the 2F(250 μs), 6W1U for the 2F(400 μs), and 6W1V for the 2F(200ms) data).

Supplementary Discussion

Water insertion mechanisms suggested in the literature: There are various hypotheses regarding the water (O_X) insertion to the Mn1 site. A direct insertion of the Ca-bound water ligand W3 (Fig. 1c, Fig. S1b) to Mn1 in a 'right-open' structure appears most straight forward, if Ca plays a role of gating substrate water intake from a water channel, as proposed earlier (50). This hypothesis is supported by a recent FTIR study (51) and theoretical studies (52-55) and consistent with our XFEL results. However, other computational studies have suggested additional options, e.g. either insertion of a Ca bound water to Mn4 or insertion of an unbound water to Mn4 or directly to Mn1. In both the 'carousel' (56) or 'pivot' (57) mechanisms, the Mn_4Ca cluster first converts from its 'open-cubane' or 'right-open' conformation (see S_2 state in Fig. 1c) to a proposed high-spin 'closed-cubane' or 'left-open' conformation in which O5 binds to Mn1 instead of Mn4, and Mn4(IV) and Mn1(III) swap oxidation state (58) (Fig. S1). The five-coordinated Mn4(III) ion may then allow insertion of either W3 or an unbound water through a shuffle of water ligands by going through the position of W1/W2 to O5 (56). The same S_3 state structure is thus formed eventually, but the history of the oxygen atoms in the various positions would be different (59). So far, however, the proposed closed-cubane high-spin S_2 state intermediate, that is necessary for the carousel or pivot mechanism, has not been observed in the room temperature crystallography experiments (12). This structure is also not preferable based on EXAFS results as the high-spin S_2 structure (60, 61). However, a recent EPR study indicates the presence of a five coordinated Mn(IV) ion in a small population of centers in the S_3 state that may be consistent with the pivot mechanism, but it is also consistent with either an open- or closed-cubane structure for the high-spin S_2 state (59). Other recent studies have proposed an open-cubane structure for the high-spin S_2 state intermediate, making the requirement for a closed-cubane structure unnecessary (62-64).

Comparison of the crystal structures with different XFEL parameters: Over the past years several studies have addressed the question of whether the choice of parameters for the fs pulses used for SFX could have an effect on the observed electronic and geometric structures. While in general no clear indication for X-ray induced changes in the determined structures and obtained transition metal X-ray spectroscopic data was found (6, 9, 20) a clear effect was demonstrated for specific conditions (65). In order to study the influence of XFEL beam parameters on the resulting spectroscopic and structural data in our case we performed experiments at both LCLS and SACLA. Here, we compare the crystal structural (Fig. S7a) and XES data (Fig. S7b) obtained at SACLA (~ 0.3 mJ/pulse at 9.5 keV, 7 fs pulse duration) and LCLS (~ 2 mJ/pulse at 9.5 keV, 37 fs pulse duration). XES $K\beta_{1,3}$ spectra of $MnCl_2$ and the 2F state of PS II matches in energy shifts and spectral shapes, showing no discernible difference using these different pulse durations and intensities. Likewise, the Fo-Fo difference density shown in Fig. S7a shows the same features both for data collected at SACLA and LCLS indicating shifts of Mn1 and Mn4, shifts of residue Glu189, presence of O_X and absence of water W20 in the 2F data versus the 0F data.

XES data of crystals and solution samples: The XES $K\beta_{1,3}$ spectra of two flash states for crystal and solution samples of PS II are shown in Fig. S1, yielding the same shift in $K\beta_{1,3}$ energies and difference spectra. The spectra are constructed from the data of 22,606 and 67,219 XFEL shots for crystal 0F and 2F, and 20,760 and 19,367 XFEL shots for solution 0F and 2F, respectively. Here, the hit rates are approximately 58% for the crystal runs, and 76% for the solution runs (see Ref. (6) for more details). The error estimates of first moments scales with the square root of the number of photons, and we note that the crystal 0F photon count is approximately half of that of either of the two solution data sets. This results in a slightly less

noisy difference spectrum (Fig. S8) and generally smaller first moment errors of the solution results (Fig. 5).

Depending on the kinetic model choosen the apparent discrepancy between the change in 1st moment and the O_X occupancy at the 50 μs time point could be explained in two ways (Fig. S9). Either it could be caused by a slightly delayed O_X insertion relative to Mn oxidation, assuming one step kinetics for the observed oxidation state change, resulting in.a slightly higher level of observed oxidation state change compared to observed O_X occupancy (Fig. S9e). Alternatively assuming a two step process with a first step that does not involve Mn oxidation state change ("lag phase") a reasonable agreement of the predicted S_3 population based on XES with the O_X density can also be obtained (Fig. S9f). Based on our current data both models, postulating 'Mn oxidation first' or 'water binding first' intermediates, are feasible and measurements from earlier time points within the $S_2 \rightarrow S_3$ transition with better signal/noise level will be required to resolve this question.

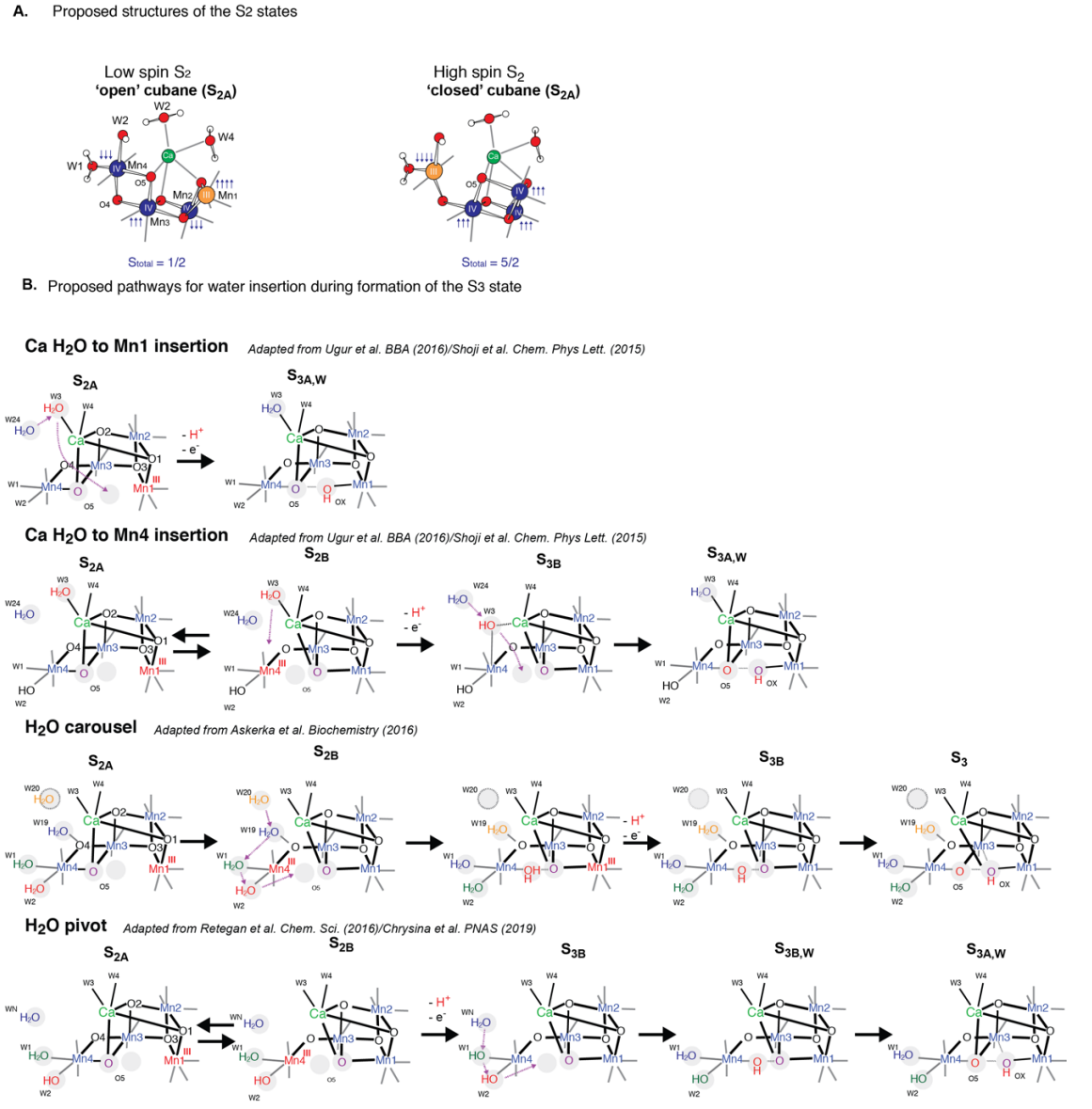

Fig. S1. Proposed structures for the S_2 state and possible pathways for water insertion during the S_3 state formation. **(a)** Proposed structures for the high spin and low spin electronic configuration of the Mn-cluster in the S_2 state are shown, with the "open" or "right open" conformation assumed for the low spin (S=1/2) state (left) and the "closed" or "left open" conformation assigned for the high spin (S=5/2) state (right). **(b)** Selection of pathways suggested in the literature for water insertion during the $S_2 \rightarrow S_3$ transition. Waters/oxygens are numbered according to our structural model and color coded to illustrate proposed water movements between the different binding sites/positions. Mn(IV) is shown in blue, Mn(III) in red. The direct insertion of an external water (not prebound at the cluster) to the Mn1 site as for example discussed in (66, 67) is not shown. Insertion of a Ca bound water either to Mn1 or to Mn4 have been suggested in several publications (see e.g. (51, 53, 68-70)). Insertion of a 2nd shell water to Mn1 was suggested in the "carousel" mechanism, described in refs (56, 71, 72), as well as in the "pivot" mechanism in (57, 59, 73). In these cases O5 binds to Mn1 instead of Mn4, and Mn4(IV) and Mn1(III) swap oxidation state. The five-coordinated Mn4(III) ion may then allow insertion of either W3 or an unbound water through a shuffle of water ligands by going through the position of W1/W2 to O5 (56). Thereafter, the open-cubane S_3 ($S_{3,A}$) state conformation is proposed to form via a flip of bonds between O5 and O_X.

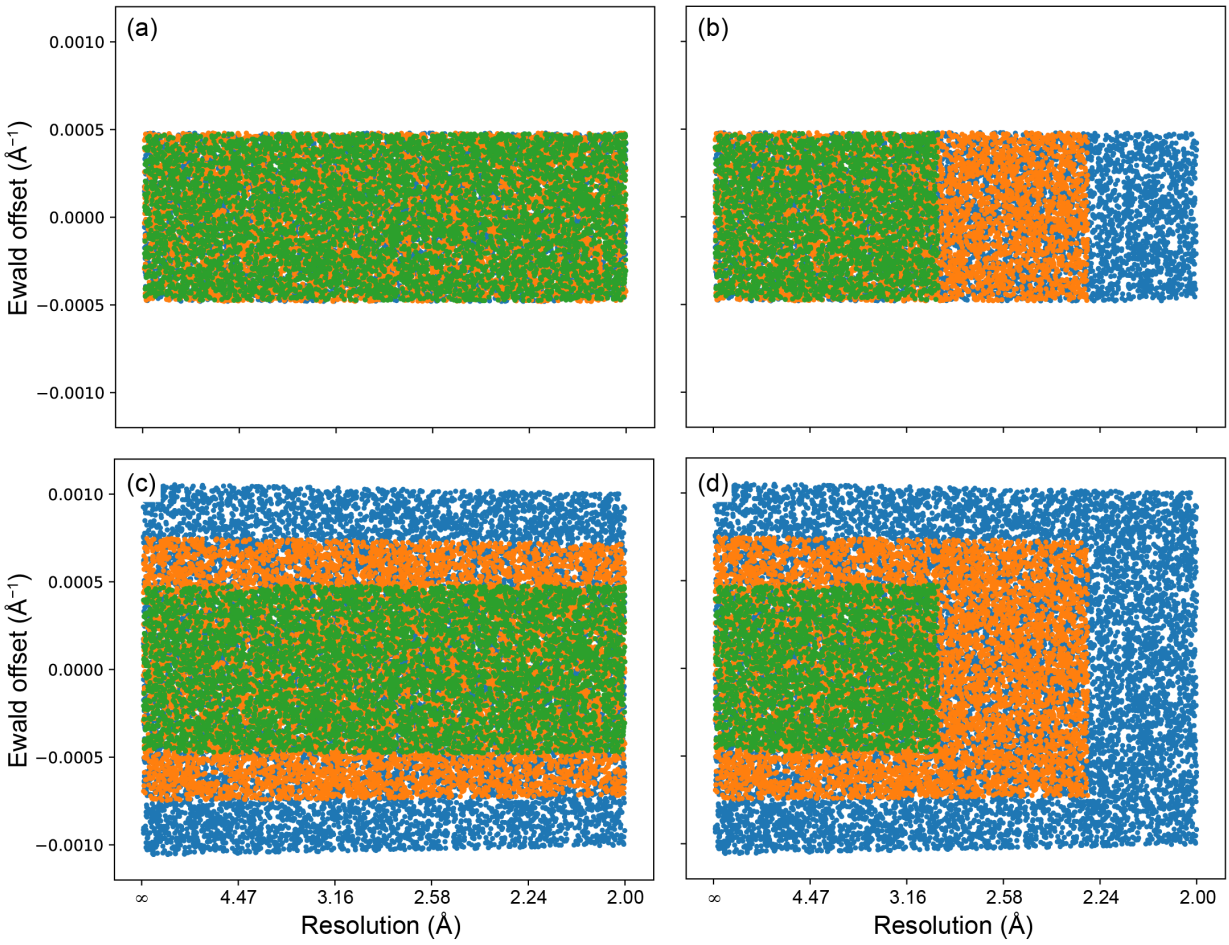

Fig. S2. Effect of parameters on the number of predicted Bragg spots. Predicted Bragg spots for XFEL data processing, sorted on axes of spatial resolution (horizontal) and Ewald offset (vertical), the latter being a measure of how far the reciprocal lattice point is from the ideal Bragg's Law diffraction condition. The plot represents three similar but differently-oriented photosystem II crystals (green, orange, and blue) under four possible scenarios. **(a)** spots from all three lattices are predicted to the same resolution (2.0 Å), and to within the same range of Ewald offset, ±0.0005 $Å^{-1}$, which corresponds to the mosaic blocks in the crystal being about 1000 Å full width (24). **(b)** spots of the orange and green lattices being predicted to progressively poorer resolutions. **(c)** spots on the orange and blue lattices predicted in successively larger Ewald offset ranges indicating smaller mosaic blocks (as reciprocal lattice point size is inversely proportional to block size). **(d)** a combination of both effects. Importantly, the number of spot predictions shown (and hence the multiplicity of coverage) is dependent only on the cutoff limits given to the integration program, and says nothing about the actual number of photons in each Bragg spot, which primarily determines data quality. For example, in panel (d), in order to prove that the larger prediction area is warranted for the blue lattice, one would have to consider the signal (or $I/\sigma(I)$ ratio) over all regions of the plot. Note, intervals on the resolution axis correspond to equal-volume shells in reciprocal space, thus the density of spots is uniform along the x-axis.

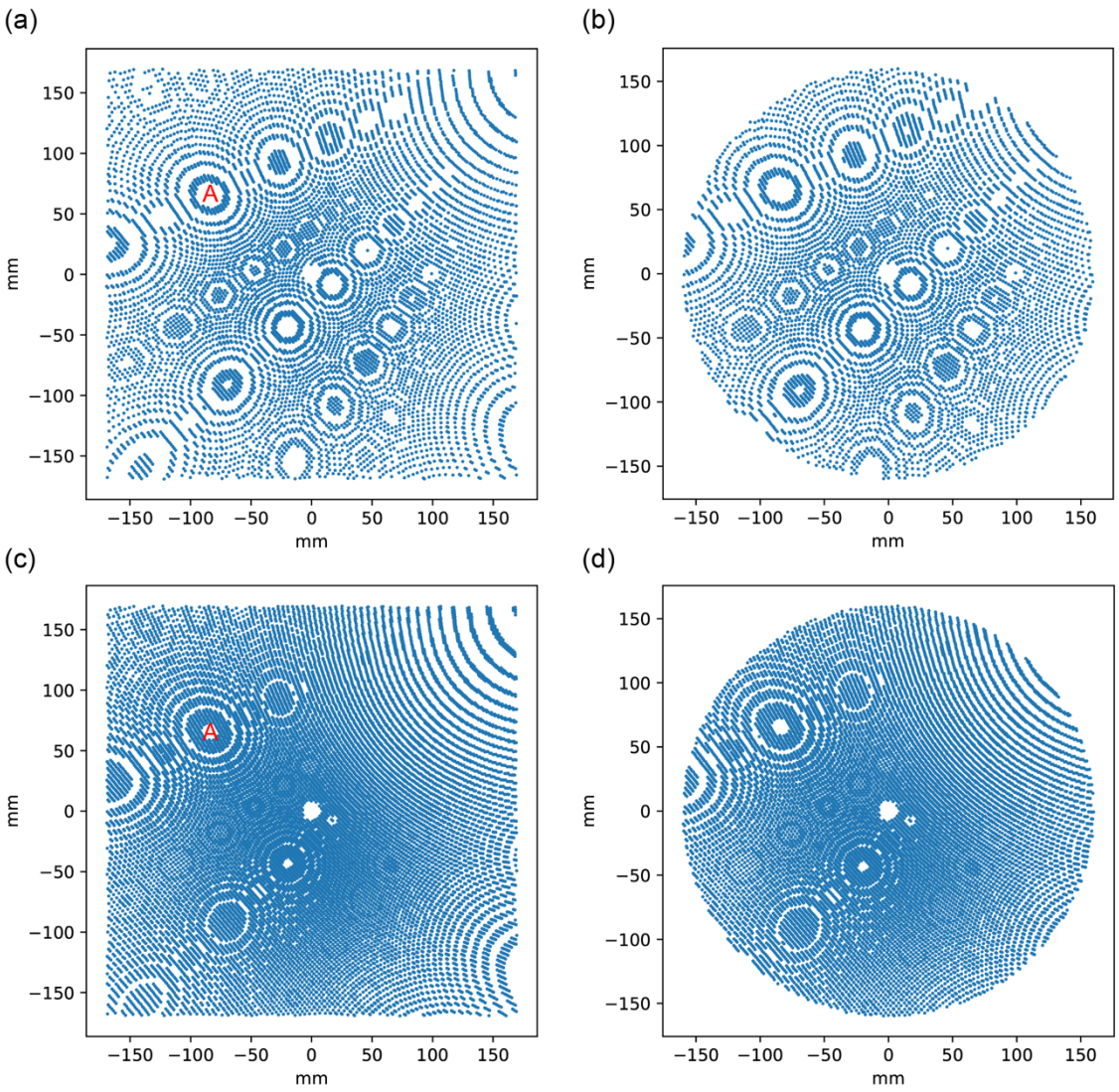

Fig. S3. Predicted Bragg spots for a single XFEL diffraction pattern from a randomly-oriented crystal of photosystem II. Four possible scenarios are considered. **(a)** spots are predicted out to the corners of the detector (7929 predictions). **(b)** spots are predicted only to an intermediate resolution, an appropriate choice when the diffraction terminates at high resolution (6351 predictions). **(c)** spots are predicted out to the corners, but on a larger range of Ewald offsets, similar to how the blue lattice is predicted in Fig. S2c. This gives rise to a thicker annulus of spots in the lunes, such as ring labeled "A", and an overall higher density of spots on the image (16953 predictions). **(d)** a similar density of spots arising from the larger range of Ewald offsets, but confined to the intermediate resolution limit (13725 predictions). As is evident, differing numbers of spot predictions can be chosen to model any diffraction pattern, but the choice of limiting resolution and Ewald offset range is only justified when measurable Bragg signal is to be found at the predicted positions. While merging statistics should differ for these cases, especially the reported multiplicity value, map quality shows little variation (Fig. S4)

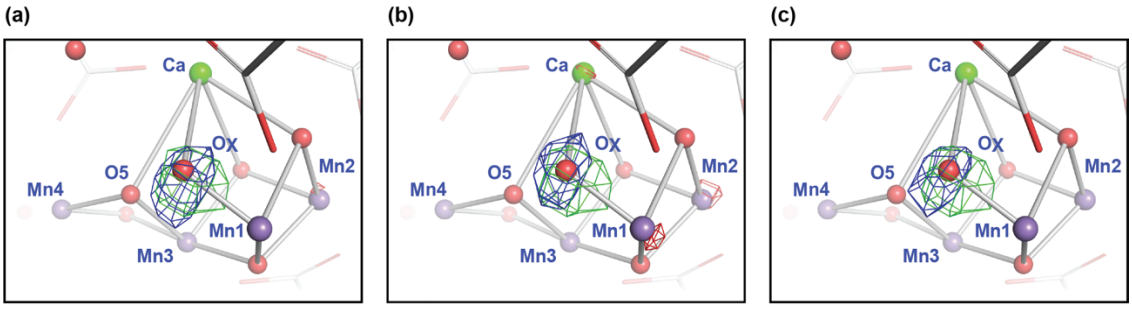

Fig S4. Omit maps (Fo-Fc) showing the region of the O_X atom in the 2F(250μs) dataset with data processed under different protocols. Maps obtained from data processed with *cctbx.xfel* using a similar (not identical) protocol as the data presented in Table 1 are shown as green/red mesh (at ±7.5 σ) together with maps based on data integrated with *CrystFEL* (blue/red mesh) with (a) all spots predicted to 2.0 Å, but varying Ewald offset cutoffs fitted to each image (±5.1 σ), or with (b) all spots predicted to 2.0 Å, but with a fixed Ewald offset cutoff applied to all images (±6.8 σ) or with (c) separate resolution cutoff fitted to each image, a fixed Ewald offset cutoff applied to all images (±5.6 σ). The central result is that the oxygen density is recovered at similar significance levels, under a variety of data processing protocols.

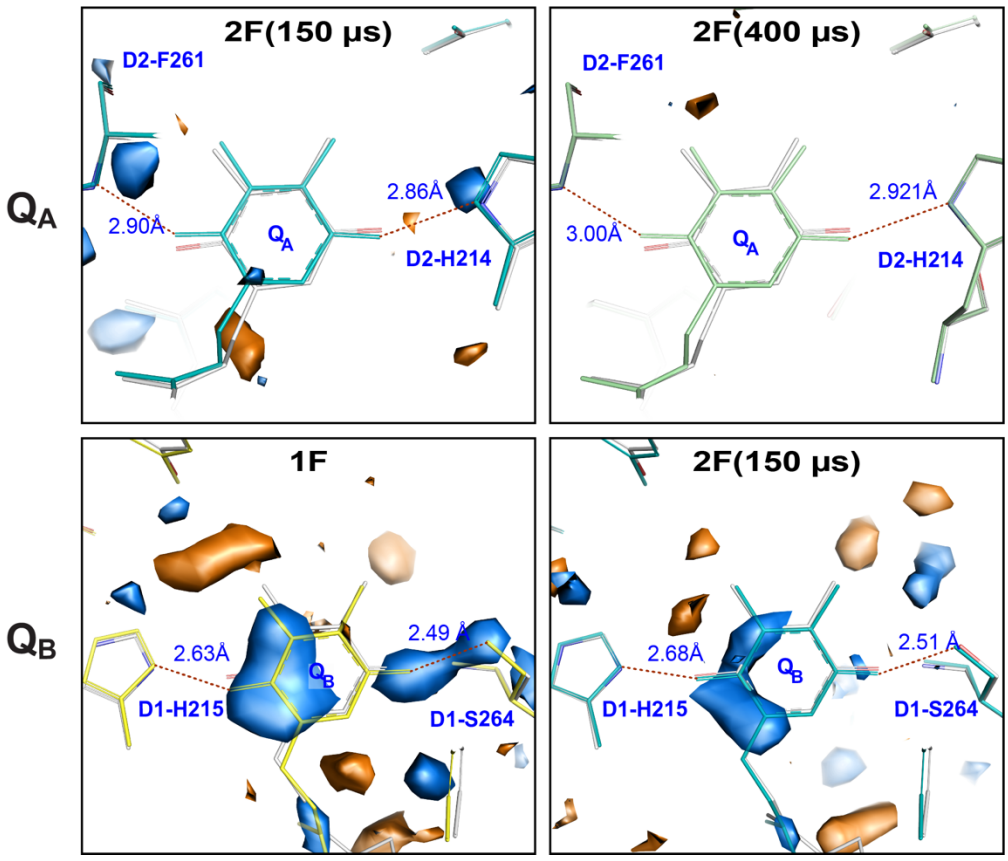

Fig. S5. Q_A/Q_B pocket changes for selected time points. Fo-Fo isomorphous difference maps (contoured at +3 σ (blue) and -3 σ (orange) (except Q_A 2F(150µs) which is contoured at 2.8σ)) of illuminated-0F diffraction data and refined models for 0F (grey) and illuminated data in the region of Q_A (top) for the 2F(150µs) and 2F(400µs) data and Q_B (bottom) for the 1F and the 2F(150µs) data.

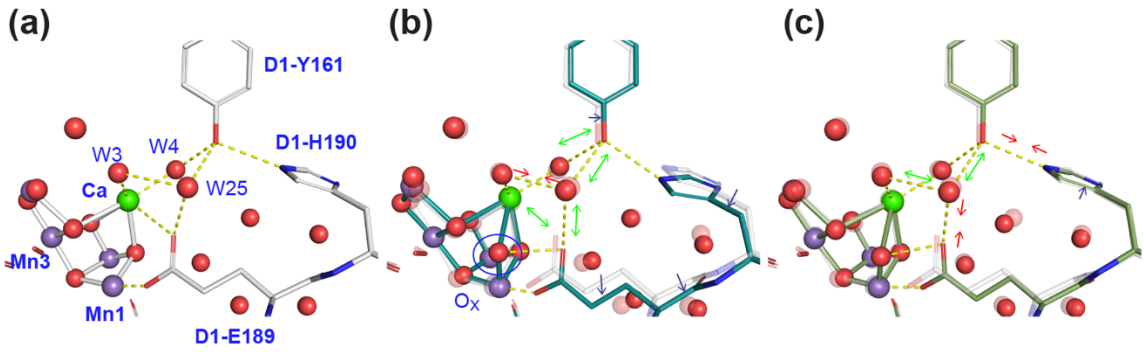

Fig. S6. The environment of Y_Z. (**a**) 0F-model (S_1) showing the Y_Z environment. Three neighboring residues are in perfect geometry for hydrogen bond interactions: His190, W4, and W25 with distances around ~2.6Å (**b**) The 2F(150μs) model (cyan) overlaid with the 0F model (light grey). The inserted O_X and its potential H-bond to Glu189 are indicated. Positional shifts of amino acid residues are highlighted by blue arrows, distance changes by green (elongation) and red (compression) arrows. (**c**) The 2F(200ms) model (olive) overlaid with the 0F model (light grey). Changes versus the 2F(150μs) positions are indicated again by blue, green and red arrows as in (b). A clear change in the H-bonding environment of W25 is visible in the 2F(150μs) model with a weaker interaction of W25 with Y_Z and Glu189 and a shorter distance to W3, indicating a possible route of proton transfer between W3 and W25. It also becomes visible in the 2F(150μs) time point model that the changes in the side chain positions of His190 and Glu189 are coupled via a shift of the protein back bone in that region (blue arrows).

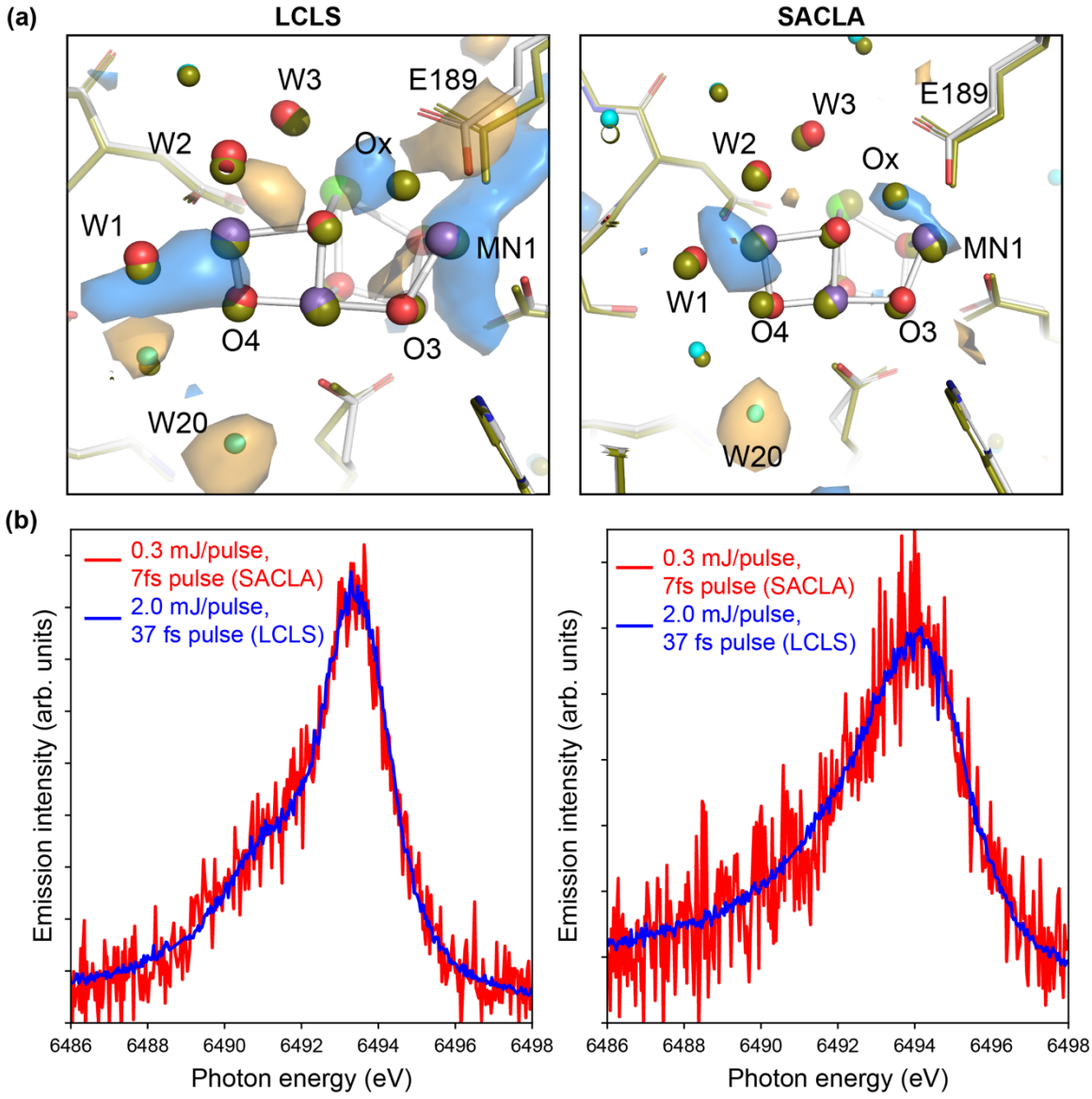

Fig. S7. Comparison of data collected at LCLS and SACLA. **(a)** The Fo-Fo maps of the 2F PS II data and **(b)** the XES K$\beta_{1,3}$ spectra of 50 mM $MnCl_2$ solution and the 2F state of PS II crystals collected at two different facilities, LCLS and SACLA. These data were collected under two different X-ray parameters; 0.3 mJ/pulse X-rays with 7 fs pulse length at SACLA, and 2.0 mJ/pulse with 37 fs pulse length at LCLS. In (b), the different noise level is due to the difference in the number of photons in a given time.

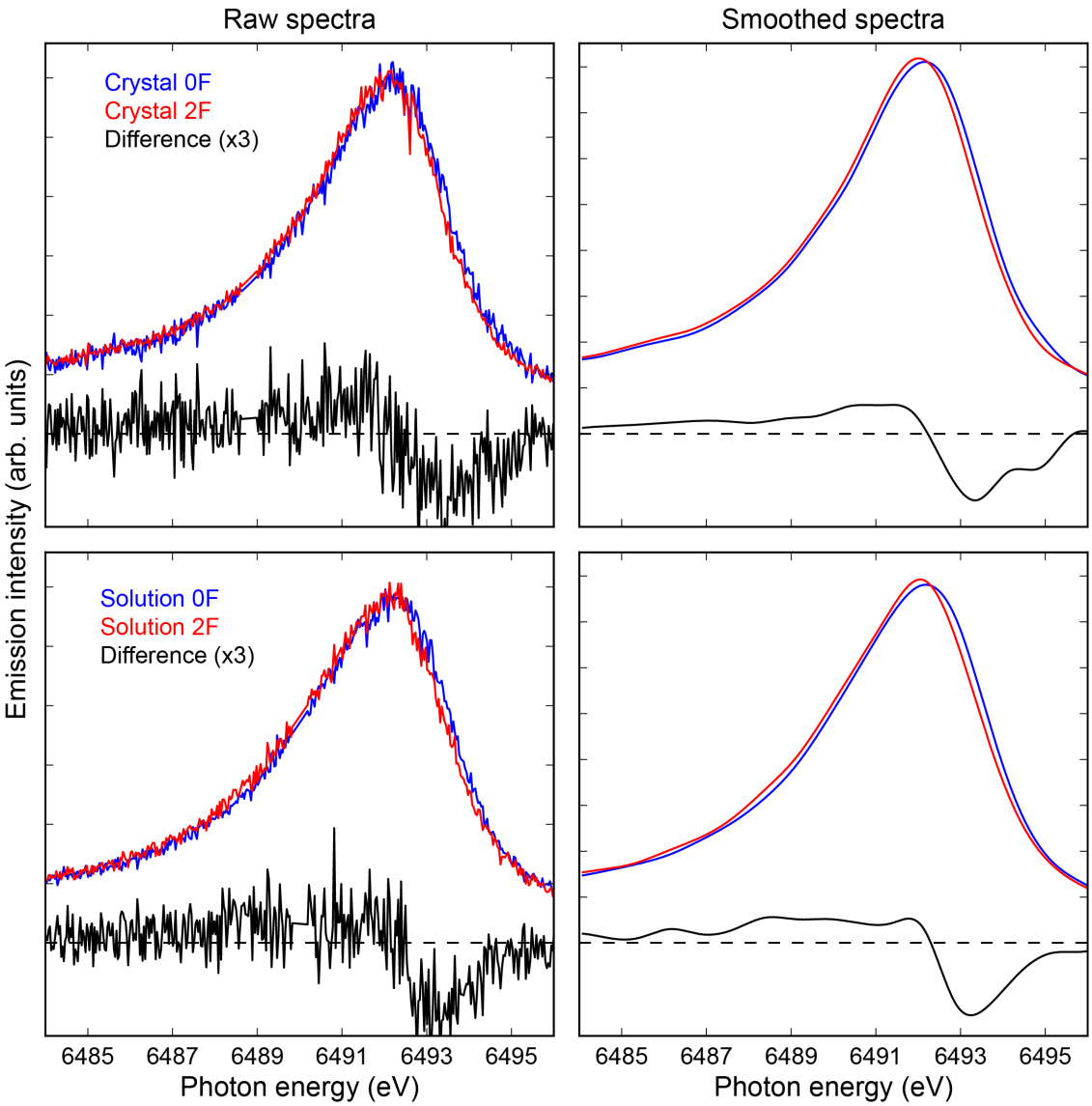

Fig. S8. Mn Kβ X-ray emission spectra collected at LCLS of 0F and 2F PSII samples for one run of microcrystals and solution, as well as corresponding difference spectra (2F-0F). Raw spectra are shown to the left and smoothed spectra to the right, where the latter are constructed by binning to a grid of 0.75 eV, followed by a cubic spline to a resolution of 0.01 eV. Spectra are area-normalized within the interval 6485 to 6497 eV.

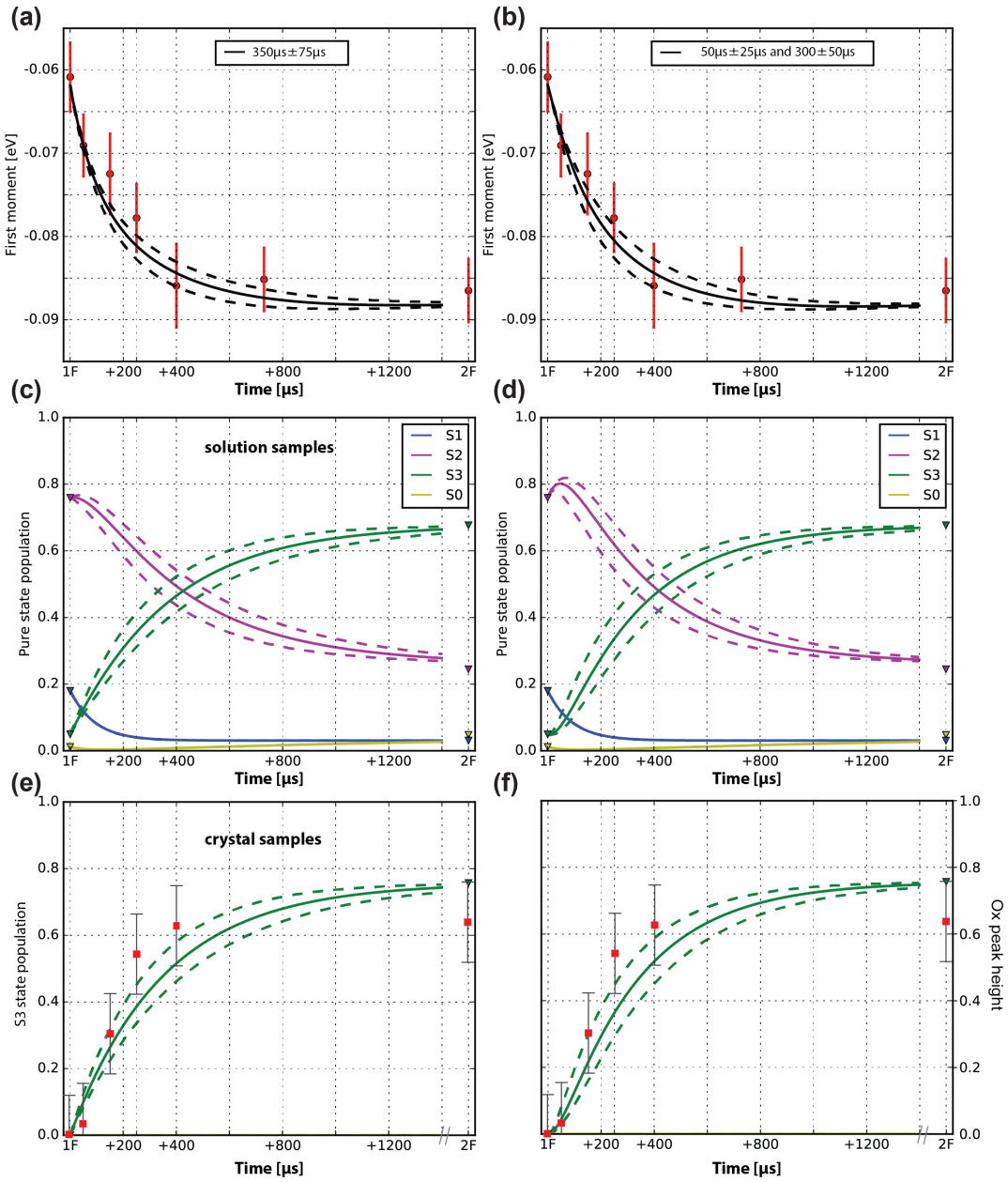

Fig. S9: XES data and kinetic traces. (a, b) Observed (non-deconvoluted) $K\beta_{1,3}$ Mn XES 1st moment changes (red circles and red error bars) after the 2nd flash are shown (as in Fig. 5c). The solid lines show the predicted combined first moment shift based on S state populations estimated for our PSII solutions samples (Table S2) and either a single time constants of τ_1=350 µs (a) or two time constants τ_1 = 50 µs and τ_2 =300 µs (b). The dashed lines illustrate the range of expected first moment shift based on an assumed error of ± 75 µs for τ_1 or ±25 µs for τ_1 and ±50 µs for τ_2.(**c,d**) Time evolution of S-state population based on the starting S-state population (Table S2) and the miss and double hit parameters for PS II solution samples and assuming a kinetic model with either one or two time constants. (**e,f**) Time evolution of the S_3-state population based on the starting S-state population and the miss and double hit parameters for PS II crystal samples and assuming a kinetic model with either one or two time constants. For comparison the O_X omit density peak height (see Fig. 5c) is plotted (red squares). Note that the O_X peak height is normalized to the O2 omit density and that due to differences in the chemical environment of O2 and O_X the value of 1.0 does not necessarily correspond to 100% occupancy of O_X.

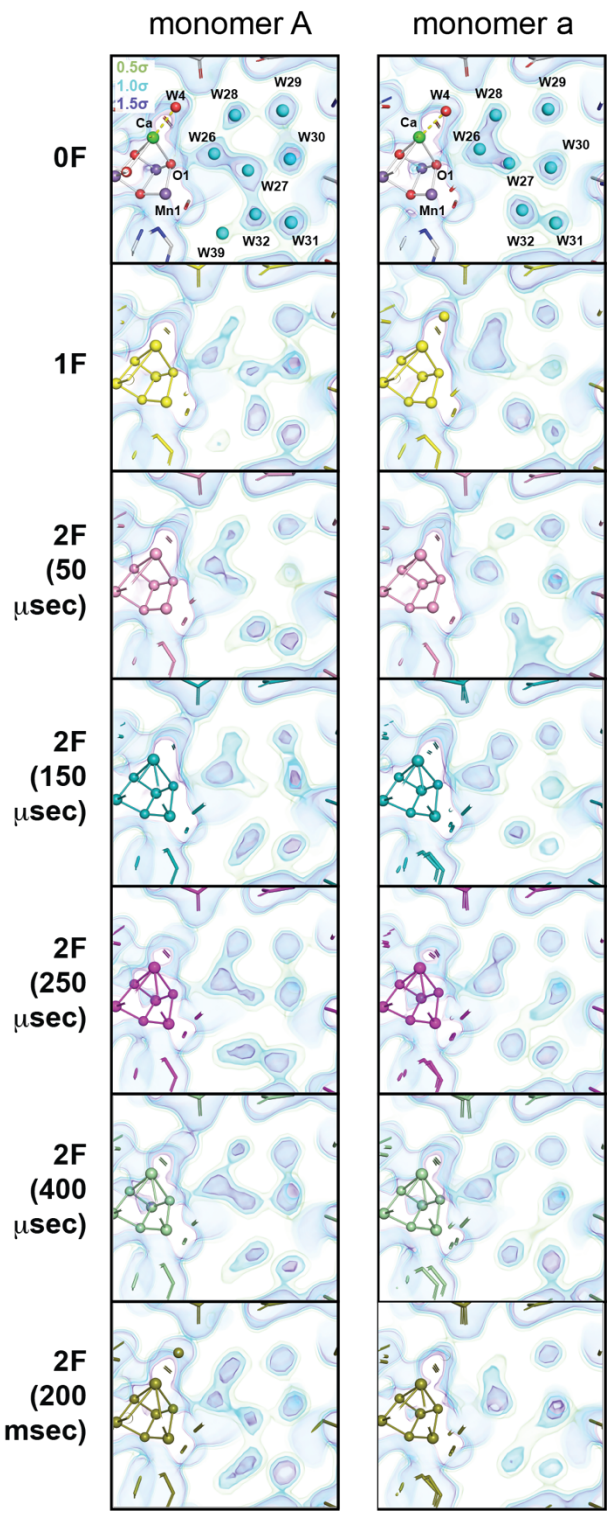

Fig. S10: Changes in the water network at the start of the O1 channel for both monomers and the different time points measured in this study. Shown is the 2Fo-Fc electron density at three different contour levels (green: 0.5σ; light blue: 1σ; blue: 1.5σ) and the refined model of the OEC and some protein residues for each of the states. For the 0F state waters are shown as cyan spheres whereas for other states waters are not shown to better visualize the changes in the electron density.

Supplementary Tables

Table S1. **Data collection, merging and refinement statistics**

	0F	1F	2F(50µs)	2F(150µs)	2F(250µs)	2F(400µs)	2F(200ms)
Resolution range refined (Å)	33.638 - 2.08	33.545 - 2.26	33.453 - 2.27	33.698 - 2.23	33.588 - 2.01	33.651 - 2.09	33.649 - 2.09
Resolution range upper bin (Å)	2.1159 - 2.08	2.299 - 2.26	2.309 - 2.27	2.2685 - 2.23	2.0345 - 2.00	2.126 - 2.09	2.216 - 2.09
Wavelength (Å)	1.302	1.302	1.302	1.302	1.303	1.302	1.302
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)	a=116.9 b=221.6 c=307.8	a=117.0 b=221.6 c=307.9	a=117.1 b=222.1 c=308.4	a=117.0 b=221.8 c=308.2	a=117.0 b=221.9 c=308.3	a=117.0 b=221.7 c=308.2	a=117.0 b=221.6 c=307.8
Lattices merged	11734	4464	5357	6195	8659	5546	10043
Unique reflections (upper bin)	474828 (23014)	370481 (17988)	367301 (17756)	386505 (18749)	535670 (25929)	468621 (22792)	468019 (22666)
Completeness (upper bin)	99.77% (97.48%)	99.60% (97.47%)	99.61% (97.06%)	99.64% (97.52%)	99.78% (97.34%)	99.72% (97.60%)	99.75% (97.42%)
$CC_{1/2}$ (upper bin)	98.0% (5.4%)	96.2% (9.1%)	96.7% (4.7%)	96.5% (5.6%)	97.7% (5.1%)	97.2% (7.7%)	97.9% (4.4%)
$I/\sigma_{Ha14}(I)^{\dagger}$ (upper bin)	13.2 (0.6)	11.8 (0.9)	11.8 (0.6)	8.0 (0.5)	14.3 (0.6)	12.3 (0.7)	12.8 (0.6)
Wilson B-factor	34.4	37.8	41.3	39.1	32.8	34.4	34.8
R-factor	18.32%	17.54%	17.98%	17.43%	17.66%	18.14%	18.44%
R-free	23.88%	23.64%	24.48%	23.36%	22.57%	23.55%	23.73%
Number of atoms	104044	103536	103319	105439	106121	106106	106350
Number non-hydrogen atoms	52408	51902	51686	52811	53496	53480	53574
Ligands	195	195	196	197	197	197	200
Waters	2143	1637	1420	1474	2159	2143	2083
Protein residues	5422	5422	5422	5422	5422	5422	5422
RMS (bonds)	0.014	0.014	0.015	0.014	0.014	0.014	0.014
RMS (angles)	1.67	1.70	1.71	1.72	1.67	1.72	1.68
Ramachandran favored	97.1%	96.80%	96.2%	96.1%	97.0%	96.8%	96.9%
Ramachandran outliers	0.1%	0.3%	0.5%	0.4%	0.3%	0.2%	0.2%
Clashscore	4.4	4.9	5.4	5.7	4.4	5.6	5.0
Average B-factor ($Å^2$)	41.5	44.2	48.5	46.4	40.0	39.9	42.4

† as defined in (28)

Table S2. The estimated S-state population (in %) for each of the flash states from fitting of the flash induced O_2 evolution pattern of a suspension of PS II crystals at pH 6.5 and solution (Sol.) samples. The S_0 population of solution samples is estimated to be 10% based on XES data. The miss parameter for crystal suspension is 13% and for solution sample is 11%. There is a 3% cross illumination due closer separation of drops on drop during collection of solution data.

	S_0		S_1		S_2		S_3	
	Crystal	Sol.	Crystal	Sol.	Crystal	Sol.	Crystal	Sol.
0F	0	10	100	90	0	0	0	0
1F	0	1	13	19	87	78	0	2
2F	0	5	2	3	23	24	75	68
3F	66	59	0	6	4	5	30	30

SI References

1. J. Hellmich *et al.*, Native-like photosystem II superstructure at 2.44 Å resolution through detergent extraction from the protein crystal. *Structure* **22**, 1607-1615 (2014).
2. M. Ibrahim *et al.*, Improvements in serial femtosecond crystallography of photosystem II by optimizing crystal uniformity using microseeding procedures. *Struct. Dyn.* **2**, 041705 (2015).
3. I. D. Young *et al.*, Structure of photosystem II and substrate binding at room temperature. *Nature* **540**, 453-457 (2016).
4. A. Guskov *et al.*, Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. *Nature Struct. Mol. Biol.* **16**, 334-342 (2009).
5. R. Krivanek, J. Kern, A. Zouni, H. Dau, M. Haumann, Spare quinones in the Q_B cavity of crystallized photosystem II of *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta Bioenerg.* **1767**, 520-527 (2007).
6. T. Fransson *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).
7. K. Beckmann, J. Messinger, M. R. Badger, T. Wydrzynski, W. Hillier, On-line mass spectrometry: membrane inlet sampling. *Photsynth. Res.* **102**, 511-522 (2009).
8. M. H. Cheah *et al.*, Assessment of the manganese cluster's oxidation state via photoactivation of photosystem II microcrystals. 10.1073/pnas.1915879117 (2019).
9. J. Kern *et al.*, Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science* **340**, 491-495 (2013).
10. F. D. Fuller *et al.*, Drop-on-demand sample delivery for studying biocatalysts in action at XFELs *Nat. Methods* **14**, 443-449 (2017).
11. J. Kern *et al.*, Taking snapshots of photosynthetic water oxidation using femtosecond X-ray diffraction and spectroscopy. *Nat. Commun.* **5**, 4371 (2014).
12. J. Kern *et al.*, Structures of the intermediates of Kok's photosynthetic water oxidation clock. *Nature* **563**, 421-425 (2018).
13. C. G. Roessler *et al.*, Acoustic injectors for drop-On-demand serial femtosecond crystallography. *Structure* **24**, 631-640 (2016).
14. P. Emma *et al.*, First lasing and operation of an Ångstrom-wavelength free-electron laser. *Nat. Photonics* **4**, 641-647 (2010).
15. R. G. Sierra *et al.*, The macromolecular femtosecond crystallography instrument at the Linac coherent light source. *J. Synchr. Rad.* **26**, 346-357 (2019).
16. K. Tono *et al.*, Beamline, experimental stations and photon beam diagnostics for the hard x-ray free electron laser of SACLA. *New J. Phys.* **15**, 083025 (2013).
17. T. Ishikawa *et al.*, A compact X-ray free-electron laser emitting in the sub-Ångstrom region. *Nat. Photonics* **6**, 540-544 (2012).
18. M. Chollet *et al.*, The X-ray Pump-Probe instrument at the Linac Coherent Light Source. *J. Synchrotron Rad.* **22**, 503-507 (2015).
19. R. Alonso-Mori *et al.*, A multicrystal wavelength dispersive X-ray spectrometer. *Rev. Sci. Instrum.* **83**, 073114 (2012).
20. R. Alonso-Mori *et al.*, Energy-dispersive X-ray emission spectroscopy using an X-ray free-electron laser in a shot-by-shot mode. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 19103-19107 (2012).
21. N. K. Sauter, XFEL diffraction: developing processing methods to optimize data quality. *J. Synchr. Rad.* **22**, 239-248 (2015).
22. G. Winter *et al.*, DIALS: implementation and evaluation of a new integration package. *Acta Cryst. D* **74**, 85-97 (2018).
23. N. K. Sauter, J. Hattne, R. W. Grosse-Kunstleve, N. Echols, New Python-based methods for data processing. *Acta Cryst. D* **69**, 1274-1282. (2013).
24. N. K. Sauter *et al.*, Improved crystal orientation and physical properties from single-shot XFEL stills. *Acta Cryst. D* **70**, 3299-3309 (2014).

25. D. G. Waterman *et al.*, Diffraction-geometry refinement in the DIALS framework. *Acta Cryst. D* **72**, 558-575 (2016).
26. A. S. Brewster *et al.*, Improving signal strength in serial crystallography with DIALS geometry refinement. *Acta Cryst. D.* **74**, 877-894 (2018).
27. A. S. Brewster, I. D. Young, A. Lyubimov, A. Bhowmick, N. K. Sauter, Processing serial crystallpgrahpic data from XFELs or synchrotrons using the *cctbx.xfel* GUI. *Comp. Cryst. Newsletter* **10**, 22-39 (2019).
28. J. Hattne *et al.*, Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nat. Methods* **11**, 545-548 (2014).
29. A. S. Brewster (2018) Scripts and documentation for Brewster et. al. 2018.
30. A. G. Leslie, Integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **55**, 1696-1702 (1999).
31. A. S. Brewster *et al.*, SAD phasing of XFEL data depends critically on the error model. **75**, 959-968 (2019).
32. P. A. Karplus, K. Diederichs, Linking crystallographic model and data quality. *Science* **336**, 1030-1033 (2012).
33. M. Suga *et al.*, Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL. *Nature* **543**, 131-135 (2017).
34. C. Kupitz *et al.*, Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature* **513**, 261-265 (2014).
35. T. A. White *et al.*, CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Cryst.* **45**, 335-341 (2012).
36. T. A. White *et al.*, Recent developments in CrystFEL. **49**, 680-689 (2016).
37. I. Zaharieva, H. Dau, M. Haumann, Sequential and coupled proton and electron transfer events in the $S_2 \rightarrow S_3$ transition of photosynthetic water oxidation revealed by time-resolved X-ray absorption spectroscopy. *Biochemistry* **55**, 6996-7004 (2016).
38. H. Sakamoto, T. Shimizu, R. Nagao, T. Noguchi, Monitoring the reaction process during the $S_2 \rightarrow S_3$ transition in photosynthetic water oxidation using time-resolved infrared spectroscopy. *J. Am. Chem. Soc.* **139**, 2022-2029 (2017).
39. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D* **66**, 213-221 (2010).
40. P. V. Afonine *et al.*, Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst. D* **68**, 352-367 (2012).
41. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Cryst. D* **66**, 486-501 (2010).
42. M. Suga *et al.*, Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99-103 (2015).
43. N. A. Law, M. T. Caudle, V. L. Pecoraro, "Manganese redox enzymes and model systems: Properties, structures, and reactivity" in Advances in Inorganic Chemistry. (1999), vol. 46, pp. 305-440.
44. K. Wieghardt, The active-Sites in manganese-containing metalloproteins and inorganic model complexes. *Angew. Chem.-Int. Edit. Engl.* **28**, 1153-1172 (1989).
45. R. M. Cinco *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).
46. S. Mukhopadhyay, S. K. Mandal, S. Bhaduri, W. H. Armstrong, Manganese clusters with relevance to photosystem II. *Chem. Rev.* **104**, 3981-4026 (2004).
47. E. Y. Tsui, J. S. Kanady, T. Agapie, Synthetic cluster models of biological and heterogeneous manganese catalysts for O_2 evolution. *Inorg. Chem.* **52**, 13833-13848 (2013).
48. Schrödinger LLC, The PyMOL molecular graphics system, version 1.8. (2015).
49. P. T. Lang, J. M. Holton, J. S. Fraser, T. Alber, Protein structural ensembles are revealed by redefining X-ray electron density noise. **111**, 237-242 (2014).
50. J. Tso, M. Sivaraja, G. C. Dismukes, Calcium limits substrate accessibility or reactivity at the manganese cluster in photosynthetic water oxidation. *Biochemistry* **30**, 4734-4739 (1991).

51. C. J. Kim, R. J. Debus, One of the substrate waters for O_2 formation in photosystem II is provided by the water-splitting Mn_4CaO_5 cluster's Ca^{2+} Ion. *Biochemistry* **58**, 3185-3192 (2019).
52. P. E. M. Siegbahn, Structures and energetics for O_2 formation in photosystem II. *Acc. Chem. Res.* **42**, 1871-1880 (2009).
53. I. Ugur, A. W. Rutherford, V. R. Kaila, Redox-coupled substrate water reorganization in the active site of photosystem II-The role of calcium in substrate water delivery. *Biochim. Biophys. Acta* **1857**, 740-748 (2016).
54. M. Shoji, H. Isobe, K. Miyagawa, K. Yamaguchi, Possibility of the right-opened Mn-oxo intermediate (R-oxo(4444)) among all nine intermediates in the S_3 state of the oxygen-evolving complex of photosystem II revealed by large-scale QM/MM calculations. *Chem. Phys.* **518**, 81-90 (2019).
55. K. Yamaguchi *et al.*, Theoretical and computational investigations of geometrical, electronic and spin structures of the $CaMn_4O_X$ (X=5, 6) cluster in the Kok cycle S_i (i=0-3) of oxygen evolving complex of photosystem II. *Physiol. Plantarum* **166**, 44-59 (2019).
56. M. Askerka, G. W. Brudvig, V. S. Batista, The O_2-evolving complex of photosystem II: Recent insights from quantum mechanics/molecular mechanics (QM/MM), extended X-ray absorption fine structure (EXAFS), and femtosecond X-ray crystallography data. *Acc. Chem. Res.* **50**, 41-48 (2017).
57. M. Retegan *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).
58. D. A. Pantazis, W. Ames, N. Cox, W. Lubitz, F. Neese, Two interconvertible structures that explain the spectroscopic properties of the oxygen-evolving complex of photosystem II in the S_2 state. *Angew. Chem. Int. Edit.* **51**, 9935-9940 (2012).
59. M. Chrysina *et al.*, Five-coordinate Mn(IV) intermediate in the activation of nature's water splitting cofactor. *Proc. Natl. Acad. Sci. USA* **116**, 16841-16846 (2019).
60. R. Chatterjee *et al.*, Structural changes correlated with magnetic spin state isomorphism in the S_2 state of the Mn_4CaO_5 cluster in the oxygen-evolving complex of photosystem II. *Chem. Sci.* **7**, 5236-5248 (2016).
61. R. Chatterjee *et al.*, Structural Isomers of the S_2 state in Photosystem II: Do they exist at room temperature and are they important for function? *Physiol. Plantarum* **166**, 60-72 (2019).
62. T. A. Corry, P. J. O'Malley, Proton isomers rationalize the high- and low-spin forms of the S_2 state intermediate in the water-oxidizing reaction of photosystem II. *J. Phys. Chem. Letts.* **10**, 5226-5230 (2019).
63. P. E. M. Siegbahn, The S_2 to S_3 transition for water oxidation in PSII (photosystem II), revisited. *Phys Chem Chem Phys* **20**, 22926-22931 (2018).
64. A. Boussac, Temperature dependence of the high-spin S_2 to S_3 transition in Photosystem II: Mechanistic consequences. *Biochim Biophys Acta Bioenerg* **1860**, 508-518 (2019).
65. K. Nass *et al.*, Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. *Journal of synchrotron radiation* **22**, 225-238 (2015).
66. P. E. M. Siegbahn, Substrate water exchange for the oxygen evolving complex in PSII in the S_1, S_2, and S_3 states. *J. Am. Chem. Soc.* **135**, 9442-9449 (2013).
67. P. E. M. Siegbahn, Water oxidation in photosystem II: oxygen release, proton release and the effect of chloride. 10063-10068 (2009).
68. C. J. Kim, H. Bao, R. L. Burnap, R. J. Debus, Impact of D1-V185 on the water molecules that facilitate O_2 formation by the catalytic Mn_4CaO_5 cluster in photosystem II. *Biochemistry* **57**, 4299-4311 (2018).
69. D. Bovi, D. Narzi, L. Guidoni, The S_2 state of the oxygen-evolving complex of photosystem II explored by QM/MM dynamics: Spin surfaces and metastable states suggest a reaction path towards the S_3 state. *Angew. Chem.-Int Edition* **52**, 11744-11749 (2013).
70. M. Shoji, H. Isobe, K. Yamaguchi, QM/MM study of the S_2 to S_3 transition reaction in the oxygen-evolving complex of photosystem II. **636**, 172-179 (2015).
71. M. Askerka, D. J. Vinyard, G. W. Brudvig, V. S. Batista, NH_3 binding to the S_2 state of the O_2-evolving complex of photosystem II: Analogue to H_2O binding during the $S_2 \rightarrow S_3$ transition. *Biochemistry* **54**, 5783-5786 (2015).

72. M. Askerka, J. M. Wang, D. J. Vinyard, G. W. Brudvig, V. S. Batista, S_3 state of the O_2-evolving complex of photosystem II: Insights from QM/MM, EXAFS, and femtosecond X-ray diffraction. *Biochemistry* **55**, 981-984 (2016).
73. M. Capone, D. Narzi, D. Bovi, L. Guidoni, Mechanism of water delivery to the active site of photosystem II along the S_2 to S_3 transition. *J. Phys. Chem. Letts.* **7**, 592-596 (2016).