Title

Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL

Michihiro Suga[1,*], Fusamichi Akita[1,2*], Michihiro Sugahara[3,*], Minoru Kubo[2,3*],Yoshiki Nakajima[1,*], Takanori Nakane[4], Keitaro Yamashita[3], Yasufumi Umena[1], Makoto Nakabayashi[1], Takahiro Yamane[1], Takamitsu Nakano[1], Mamoru Suzuki[3,5], Tetsuya Masuda[3,6], Shigeyuki Inoue[3,7], Tetsunari Kimura[3,8], Takashi Nomura[3], Shinichiro Yonekura[1], Long-Jiang Yu[1], Tomohiro Sakamoto[1], Taiki Motomura[1,9], Jing-Hua Chen[1,10], Yuki Kato[11], Takumi Noguchi[11], Kensuke Tono[12], Yasumasa Joti[12,] Takashi Kameshima[12], Takaki Hatsui[3], Eriko Nango[3,13], Rie Tanaka[3], Hisashi Naitow[3], Yoshinori Matsuura[3], Ayumi Yamashita[3], Masaki Yamamoto[3], Osamu Nureki[4], Makina Yabashi[3, 12], Tetsuya, Ishikawa[3], So Iwata[3, 13,**], Jian-Ren Shen[1,9,10,**]

[1]Research Institute for Interdisciplinary Science and Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima Naka, Okayama 700-8530, Japan.

[2]Japan Science and Technology Agency, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan.

[3]RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan.

[4]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan.

[5]Institute for Protein Research, Osaka University, Yamadaoka, Suita, Osaka 565-0871, Japan.

[6]Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan.

[7]Department of Cell Biology and Anatomy, Graduate School of Medicine, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan.

[8]Department of Chemistry, Graduate School of Science, Kobe University, 1-1 Rokkodai, Nada- ku, Kobe 657-8501, Japan.
[9]Department of Picobiology, Graduate School of Life Science, University of Hyogo, 3-2-1 Kouto. Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan.
[10]Key laboratory of photobiology, Institute of Botany, Chinese Academy of Sciences, No.20 Nanxincun, Xiangshan Beijing 100093, China.
[11]Division of Material Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan.
[12]Japan Synchrotron Radiation Research Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5198, Japan.
[13]Department of Cell Biology, Graduate School of Medicine, Kyoto University, Yoshidakonoe-cho, Sakyo-ku, Kyoto, 606-8501, Japan.

*These authors contributed equally to this work.

**Corresponding authors.

Summary

Photosystem II (PSII) is a huge membrane-protein complex consisting of 20 different subunits with a total molecular mass of 350 kDa for a monomer, and catalyzes light-driven water oxidation at its catalytic center, the oxygen-evolving complex (OEC)[1-3]. The structure of PSII has been analyzed at 1.9 Å resolution by synchrotron radiation X-rays, which revealed that OEC is a Mn_4CaO_5 cluster organized in an asymmetric, "distorted-chair" form[4]. This structure was further analyzed with femtosecond X-ray free electron lasers (XFEL), providing the “radiation damage-free”[5] structure. The mechanism of O=O bond formation, however, remains obscure due to the lack of intermediate state structures. Here we report the structural changes of PSII induced by 2-flash (2F) illumination at room temperature at a resolution of 2.35 Å using time-resolved serial femtosecond crystallography (TR-SFX) with an XFEL provided by the SPring-8 angstrom compact free-electron laser (SACLA). Isomorphous difference Fourier map between the 2F and dark-adapted states revealed two areas of apparent changes; they are around Q_B/non-heme iron and the Mn_4CaO_5 cluster. The changes around the Q_B/non-heme iron region reflected the electron and proton transfers induced by the 2F-illumination. In the region around the Mn_4CaO_5 cluster, a water molecule located 3.5 Å from the Mn_4CaO_5 cluster disappeared from the map upon 2F illumination, leading to a closer distance between another water molecule and O4, suggesting also the occurrence of proton transfer. Importantly, the 2F-dark isomorphous difference Fourier map showed an apparent positive peak around O5, a unique μ_3-oxo-bridge located in the quasi-center of Mn1 and Mn4[4,5]. This suggests an insertion of a new oxygen atom (O6) close to O5, providing an O=O distance of 1.5 Å between these two oxygen atoms. This provides a mechanism for the O=O bond formation

consistent with that proposed by Siegbahn[6,7].

Main text

Fig. 1a shows organization of the electron transfer chain of PSII in a pseudo-C2 symmetry by two subunits D1 and D2. The water-oxidation reaction proceeds via the S*i*-state cycle[8] (with *i*=0-4), where dioxygen is produced in the transition of $S_3 \rightarrow (S_4) \rightarrow S_0$ (Fig. 1b). The high-resolution structures of PSII analyzed so far were for the dark-stable S_1 state[4,5], although a few studies on the low-resolution intermediate S-state structures have been reported by TR-SFX[9-11]. During the revision of our manuscript, Young et al. reported a 2F-illuminated state structure at 2.25 Å resolution where no apparent changes around O5 were observed[12], although estimations of the resolution could yield somewhat different values so that small movement of some water molecules may escape the detection. In order to achieve resolution high enough to uncover small structural changes induced by flash illuminations yet allowing *Si*-state transition to proceed efficiently, we determined the optimal crystal size of PSII with a maximum length of 100 μm, which diffracted up to a resolution of 2.1 Å by a SACLA-XFEL pulse (Extended Data Fig. 1, Methods). We illuminated the small crystals by 2F (see Methods and Fig. 1), which gave rise to a population of 46% S_3 estimated by Fourier transform infrared (FTIR) difference spectroscopy (Extended Data Fig. 2, Methods).

The PSII microcrystals were mixed with a grease matrix and delivered to the XFEL beam by a high viscosity micro-extrusion injector[13], which greatly reduced the sample consumption and provided a stable flow for the uniform excitation. To ensure the turnover of PSII, potassium ferricyanide was added into the samples, and the samples

were pre-flashed before the measurements for synchronization of PSII into the S_1 state and the oxidation of Y_D. Diffraction data were also collected from samples without pre-flash (non-pre-flash), and they will be discussed where differences were found with the pre-flashed samples.

Two datasets were collected for each of the samples (either with or without pre-flashes), one is from the dark-state and the other one is from 2F-illuminated samples. The two datasets for the pre-flashed samples were analyzed to 2.35 Å resolution, and those for the non-pre-flashed samples were analyzed to 2.50 Å resolution (Extended Data Fig. 1, Table 1 and 2); all of the datasets had very high multiplicities (see Methods). Fig. 1e and f showed part of the electron density map from the non-pre-flashed dark-sample (similar maps were obtained with pre-flashed samples, and we will focus on the results of pre-flashed samples in the following unless otherwise described), where the electron densities for the orientations of amino acid side chains and water molecules are clearly visible. We were able to build 1,626 waters for the dark-state and 1,584 waters for the 2F-state, respectively. The manganese atoms and calcium atom in OEC were distinguished even in the anomalous difference Fourier map with peak intensities well beyond the noise level, owing to the long wavelength of the XFEL used. For example, among top 30 peaks in the anomalous map of the non-pre-flashed 2F-data (5GTI), top 7 peaks with intensities above 15.0σ were identified as the manganese atoms, and other 23 peaks were identified as the calcium (including the newly identified calcium site at the stromal side of CP43), chloride or sulfur atoms (Extended Data Fig. 3 and Extended Data Table 3). This indicated the accurate measurement and integration of the diffraction intensities. Nevertheless, the interatomic distances within OEC could not be refined accurately because of the limited resolution. Thus, tight Mn-O distance

restraints derived from the S_1 state 1.95 Å structure[5] were used during the refinement of the dark data and relatively loose restraints were applied for the 2F-data (Methods). Therefore we will focus on the structural changes induced by the two flashes based on the isomorphous difference Fourier map between the two datasets.

The two datasets from the two different states were highly isomorphous with a R_{iso} of 0.068 at 2.35 Å resolution, enabling the approach of isomorphous difference Fourier map possible (Extended Data Table 1). Superposition of the two state structures yielded a root mean square deviation of 0.10 Å. The corresponding R_{iso} values between the pre-flashed and non-pre-flashed samples were more than 0.118 at 2.5 Å resolution, which may be caused by the different batches of samples used or differences in the post-crystallization processes employed.

Based on the isomorphous difference Fourier map between the 2F- and dark-states [F(2F) – F(dark) difference map], two areas with major light-induced structural changes were observed (Fig. 2). These areas are localized around the Q_B/non-heme iron-binding site and OEC (areas cycled by dotted red lines in Fig. 2b), and similar changes were observed in both monomers of the dimer (a third area with apparent difference Fourier peaks was found in the vicinity of Y_D in non-pre-flashed samples, see below and Extended Data Fig. 4). All of these areas are related to the electron and proton transfer in PSII, suggesting the efficient laser pulse excitation. The peaks were detected at 8.6σ for the Q_B-site, 10.7σ for OEC in monomer-A; and 6.9σ for the Q_B-site, 12.7σ for OEC in monomer-B. The largest Fourier-difference peak in the other areas of PSII was found to be 4.4σ, indicating that the peaks observed in these two areas are well above the noise level.

Q_B forms H-bonds with D1-His215 and D1-Ser264; the former is a ligand for the

non-heme iron. The isoprenoid tail of Q_B is located in a hydrophobic environment formed by D1-Phe211, Met214, Phe255, Phe265, Phe274, $Pheo_{D2}$ and lipids (Fig. 3a). In the high resolution S_1 state structures[4,5], Q_B exhibited a weak electron density with a high B-factor than Q_A possibly due to the existence of its mixed protonation states. After 2F illumination, positive electron densities were observed in its head group region whereas a pair of positive and negative densities was observed in the isoprenoid tail in the Fourier difference map, and the average temperature factor of the quinone head group was decreased from 116 $Å^2$ to 91 $Å^2$. This indicates a rotation of its head group around the axis vertical to the benzene ring by about 10 degrees toward D1-Ser264 and D1-His215, which shortens the distances between the head group of Q_B and these two residues by 0.1-0.2 Å and therefore strengthens the H-bonds between them. This can be explained as follows. Since we added potassium ferricyanide to the crystals well before measurements, the non-heme iron is pre-oxidized to Fe^{3+}, enabling it to be able to accept one electron upon flash illumination. Thus, the 2F illumination resulted in a dominant species of Q_B^-, which will make stronger H-bonds with Ser264 and His215, resulting in the rotation of the head group and the decrease of the temperature factor. In relation to this, the hydrophobic residues D1-Phe211, Phe255, Phe265 and Met214 surrounding the isoprenoid tail also moved slightly to adjust the cavity of the Q_B-site (Fig. 3a-b). Interestingly, a loop region involving D1-Asn266-Ser268, which shields the cavity of the Q_B-site in the dark structure, moved by up to 0.8 Å upon 2F illumination, resulting in the partial opening of the Q_B-site which may suggest a route for the quinone exchange.

In the vicinity of Q_B, a strong negative peak was observed between the bicarbonate (BCT) and the non-heme iron in the difference Fourier map. There was a smaller

positive peak in the opposite side of BCT, giving rise to a positional shift of BCT up to 0.5 Å and an elongation of distances between BCT and the non-heme iron up to 0.2 Å (Fig 3b). This may be due to a rearrangement of the H-bond network around BCT and/or a reorientation of BCT upon 2F illumination, which may be caused by the involvement of BCT in the proton-transfer for Q_B protonation or the reduction of the non-heme iron. The possible involvement of BCT in the proton transfer is in agreement with the observation that removal of BCT resulted in a retardation of the quinone exchange reaction[14]. In contrast to the Q_B-site, no large changes were found in the Q_A-site.

A number of Fourier difference peaks were found in the region around OEC (Fig. 3c, d), based on which, the following major structural changes can be identified: (i) Mn4 moved slightly toward outside of the cubane, resulting in a slight elongation in its distance with Mn1 by 0.1-0.2 Å; (ii) Ca^{2+} also moved away from Mn4 slightly; (iii) A new, strong positive peak was found close to O5, which was modeled as a new oxygen atom "O6"; (iv) The side chain of D1-Glu189 moved away from the cubane to accommodate O6; (v) A strong negative peak was found on the water molecule W665, indicating that this water molecule was displaced from this position upon 2F-illumination. In relation to this, the water molecule W567 hydrogen-bonded to O4 moved toward O4, resulting in a closer distance between these two oxygen species; (vi) Accompanying these changes, some other ligand residues also exhibited slight structural changes; these include D1-Asp61, Asp170, His332 and Ala344.

Among the structural changes observed above, the displacement of W665 and the appearance of a new oxygen atom O6 provided important implications for the mechanism of water oxidation. Due to the displacement of W665, the hydrogen-bond

between this water and W567 was broken. This caused the movement of W567 toward O4, as they are already H-bonded (Fig. 3c). The average distance between W567 and O4 upon 2F-illumination was determined to be 2.32 Å from the two monomers. This distance is extremely short, which may suggest that W567 is a hydroxide ion rather than a water in the S_3 state. This raises a possibility that these two oxygen species represent a state prepared for the upcoming O=O bond formation (Fig. 4). Indeed, studies with W-band ^{17}O electron-electron double resonance-detected NMR has suggested that either O4 or O5 is exchangeable and thus may serve as candidate for the substrate water[15]. An earlier theoretical study also suggested that O4 or O5 may be one of the substrates for O=O bond formation[16]. However, if we consider the possible experimental errors in the inter-atomic distances determined at the current resolution, these two oxygen atoms may still be connected by a H-bond instead of forming a pre-state for the O=O bond formation. The structural changes observed therefore strongly suggest the occurrence of proton transfer around this region, in agreement with the recent reports that the H-bond network starting from O4, W567 and W665 (Extended Data Fig. 5) may serve as a path for protons released from OEC in the transition from S_0 to S_1[17.18]. The movement of W665 observed in the present study may be related to this proton transfer. Based on the proton release pattern of 1:0:1:2 for the $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4$ transitions, however, the proton transfer from W567-W665 may occur in the S_2 to S_3 transition.

The O6 atom in the 2F-illuminated state has a very close distance to O5 (average distance of 1.5 Å for the two monomers) (Fig. 3c-f). This enables them to form a peroxide or superoxide, etc[19]. This strongly suggests that this site is the reaction site for the O=O bond formation (Fig. 4), which is in agreement with the model proposed by Siegbahn et al[6,7,20] as well as the results of electron paramagnetic resonance

measurements[21,22] and quantum mechanics/molecular mechanics (QM/MM) calculations[23]. The remaining structural changes can be easily explained by the insertion of this oxygen atom: The elongation of the distance between Mn4 and Mn1 is apparently due to the requirement to accommodate this new oxygen atom. In particular, the movement of D1-Glu189 away from the cubane is necessary to accommodate O6, as this results in an average distance of 2.70 Å between O6 and Glu189-OE2. Without this movement, the distance between O6 and Glu189-OE2 would be 2.30 Å, which will hinder the insertion of this oxygen atom. This is made possible by the unique feature of Glu189 as a monodentate ligand to OEC among all of the residues. We should also point out that, O6 is not necessarily a new water molecule from the outside of PSII, as some water molecule already present in the S_1 state may move to this position[4,19,24].

In addition to the above two major areas of changes with the pre-flashed samples, a third area with structural changes was found around Y_D (D2-Tyr160) in the non-pre-flashed samples. Y_D is partially H-bonded to a water molecule W508 which exhibits two conformations, one H-bonded (W508I) and the other one not H-bonded (W508II) to Y_D. In the Fourier difference map of the non-pre-flashed samples, a negative peak was found in the position of W580I, suggesting its displacement, resulting in the breakage of the H-bond with Y_D (Extended Data Fig. 4a-c). This suggests the proton transfer around this region upon 2F-illumination, a finding in agreement with the recent QM/MM calculations where the proximal W508I was found to be unstable upon oxidation of Y_D[25]. This in turn suggests partial oxidation of Y_D upon 2F illumination probably due to an inefficient electron donation from the Mn_4CaO_5-cluster in the crystals. The oxidation of Y_D will result in the proton transfer from Y_D to the bulk solvent, leading to the breakage of the H-bond between Y_D and

W508I and the movement of this water to the distal position or bulk solvent[25,26]. Thus, this structural change is again caused by the proton transfer around Y_D.

References

1. Shen, J.-R. The structure of photosystem II and the mechanism of water oxidation in photosynthesis. *Annu. Rev. Plant Biol.* **66,** 23–48 (2015).
2. Yano, J. & Yachandra, V. K. Mn_4Ca cluster in photosynthesis: where and how water is oxidized to dioxygen. *Chem. Rev.* **114,** 4175–205 (2014).
3. Najafpour, M. et al., Manganese compounds as water-oxidizing catalysts: From the natural water-oxidizing complex to nano-sized manganese oxide structures. *Chem. Rev.* **116**, 2886–2936 (2016).
4. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 A. *Nature* **473**, 55–60 (2011).
5. Suga, M. *et al.* Native structure of photosystem II at 1.95 angstrom resolution viewed by femtosecond X-ray pulses. *Nature* **517**, 99–103 (2015).
6. Siegbahn P. E. A structure-consistent mechanism for dioxygen formation in photosystem II. *Chem. Eur. J.* **14,** 8290-8302 (2008).
7. Siegbahn P. E. Structures and energetics for O_2 formation in photosystem II. *Acc. Chem. Res*. **42,** 1871-1880 (2009).
8. Kok. B., Forbush, B. & McGloin, M. Cooperation of charges in photosynthetic O_2 evolution-I. A linear four step mechanism. *Photochem Photobiol.* **11**, 457–475 (1970).
9. Kern, J. *et al.* Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science* **340**, 491–495 (2013).
10. Kupitz, C. *et al.* Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature* **513**, 261–265 (2014).
11. Kern, J. *et al.* Taking snapshots of photosynthetic water oxidation using femtosecond X-ray diffraction and spectroscopy. *Nature Commun.* **5**, 4371, doi:10.1038/ncomms5371 (2014).
12. Young, I.D. *et al.* Structure of photosystem II and substrate binding at room temperature. *Nature* **540,** 453–457 (2016).
13. Sugahara, M. *et al.* Grease matrix as a versatile carrier of proteins for serial crystallography. *Nature Methods* **12**, 61–63 (2015).
14. Shevela, D., Eaton-Rye, J. J., Shen, J. R. & Govindjee. Photosystem II and the

unique role of bicarbonate: a historical perspective. *Biochim. Biophys. Acta* **1817**, 1134–1151 (2012).

15. Rapatskiy, L. *et al.* Detection of the water-binding sites of the oxygen-evolving complex of Photosystem II using W-band ^{17}O electron-electron double resonance-detected NMR spectroscopy. *J. Am. Chem. Soc.* **134**, 16619–16634 (2012).
16. Yamanaka, S. *et al.* Possible mechanisms for the O-O bond formation in oxygen evolution reaction at the CaMn4O5(H2O)(4) cluster of PSII refined to 1.9 angstrom X-ray resolution. *Chem. Phys. Lett.* **511**, 138–145 (2011).
17. Saito, K., Rutherford, A. W. & Ishikita, H. Energetics of proton release on the first oxidation step in the water-oxidizing enzyme. *Nature Commun.* **6**, 8488, doi:10.1038/ncomms9488 (2015).
18. Takaoka, T., Sakashita, N., Saito, K. & Ishikita, H. pKa of a Proton-Conducting Water Chain in Photosystem II. *J. Phys. Chem. Lett.* **7**, 1925–1932 (2016).
19. Isobe, H. *et al*, Chemical equilibrium models for the S_3 state of the oxygen-evolving complex of photosystem II. *Inorg. Chem*. **55**, 502-511 (2016).
20. Siegbahn, P. E. Water oxidation mechanism in photosystem II, including oxidations, proton release pathways, O-O bond formation and O2 release. *Biochim. Biophys. Acta* **1827**, 1003–1019 (2013).
21. Cox, N. *et al.* Electronic structure of the oxygen-evolving complex in photosystem II prior to O-O bond formation. *Science* **345**, 804–808 (2014).
22. Perez-Navarro, M., Neese, F., Lubitz, W., Pantazis, D. A. & Cox, N. Recent developments in biological water oxidation. *Curr. Opin. Chem. Biol.* **31**, 113–119 (2016).
23. Shoji, M, Isobe, H. & Yamaguchi, K. QM/MM study of the S_2 to S_3 transition reaction in the oxygen-evolving complex of photosystem II. *Chem. Phys. Lett.* **636,** 172-179 (2015).
24. Isobe, H, *et al*, Theoretical illumination of water-inserted structures of the $CaMn_4O_5$ cluster in the S_2 and S_3 states of oxygen-evolving complex of photosystem II: full geometry optimizations by B3LYP hybrid density functional, *Dalton Trans.*, **41**, 13727-13740 (2012).
25. Saito, K., Rutherford, A. W. & Ishikita, H. Mechanism of tyrosine D oxidation in Photosystem II. *Proc. Natl Acad. Sci. USA* **110**, 7690–7695 (2013).
26. Nakamura, S. & Noguchi T. Infrared detection of a proton released from tyrosine Y_D to the bulk upon its photo-oxidation in photosystem II. *Biochemistry* **54,** 5045−5053 (2015).

Acknowledgements

We thank Keisuke Kawakami and Nobuo Kamiya for sharing unpublished information, Hiroyuki Mino for information on the flash illumination conditions, Hideo Ago and Kizashi Yamaguchi for discussions. This work was supported by a program for promoting the enhancement of research universities at Okayama University, JSPS KAKENHI Grant Nos. JP15H01642, JP16H06162, JP16H06296 (M. Suga), JP16K21181 (F. A.), JP15H05588 (Y. U.) and JP24000018 (J.-R. S.), an X-ray Free Electron Laser Priority Strategy Program (J.-R.S., S. Iwata) from MEXT, Japan, an Asahi Glass Foundation (F. A.), a Kato Memorial Bioscience Foundation (F. A.), an Inamori foundation (M. Suga), the Research Acceleration Program from Japan Science and Technology agency (JST) (S. Iwata), PRESTO from JST (M. K. and F. A.), a grant from Pioneering Project "Dynamic Structural Biology" of RIKEN (M. K.), and the Strategic Priority Research Program of CAS (XDB17030100) (J.-R. S.). The XFEL experiments were performed at beamline 3 of SACLA with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (proposal nos. 2013B1259, 2014A1243, 2014A6927, 2014B1281, 2014B6927, 2014B8048, 2015A1108, 2015A6522, 2015B2108, 2015B6522, 2015B8044, 2016A2542, 2016A6621 and 2016A8033), and we thank staffs at SACLA for their help. We also acknowledge the computational support from the SACLA HPC system and Mini-K supercomputer system.

Author Contributions

J.-R. S. conceived the project; M. Suga, F. A., M. Sugahara, M. K., Y. U. and J.-R. S. contributed to the design of the experiment; F. A., Y. N., M. N. and T. Nakano grew the

cells, purified the PSII samples; F. A., and Y. N. prepared the PSII crystals; M. Sugahara, E. N., and R. T. developed the sample delivery system; F. A., Y. N., M. N., Y. U. and M. Sugahara tested and optimized buffer and crystal suspension conditions for injection; M. Suzuki, T. Masuda, S. Inoue, T. Motomura, J.-H. C., H. N., Y. M., and A. Y. operated the injector; K. T., Y. J., T. Kameshima, T. H., M. Yabashi, T. I. and S. Iwata developed the diffraction instrumentation; M. K., T. Kimura. and T. Nomura designed and optimized the laser excitation scheme and aligned the lasers; M. Suga, F. A., M. Sugahara, Y. N., T. Nakane, K. Y., M. N., Y. U., M. Suzuki., T. Masuda, S. Inoue, S. Y., L.-J. Y., T. Motomura, J.-H. C., R. T., H. N., Y. M., A. Y. and J.-R. S. participated in collection of the X-ray diffraction data at SACLA; T. Nakane, K. Y., M. Yamamoto, O. N. and S. Iwata. developed the data evaluation and/or hit finding programs; Y. Kato, F. A. and T. N. performed FTIR analysis; M. Suga, T. Y. and T. S. analyzed the femtosecond crystallography X-ray diffraction data; M. Suga refined the structure, calculated the electron density maps and made the figures; M. Suga and J.-R. S. wrote the manuscript, and all the authors participated in the discussion of the results and writing of the paper.

Author Information

The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S. Iwata. (s.iwata@mfour.med.kyoto-u.ac.jp) or J.-R. S. (shen@cc.okayama-u.ac.jp).

Figure Legends

Figure 1 | TR-SFX of PSII at SACLA-XFEL. a, Organization of the electron transfer chain of PSII. **b,** S*i*-state cycle of the water-oxidation reaction of OEC. **c,** Experimental setup for the TR-SFX experiments of PSII. **d,** Schematic representation of the timings of two flash illuminations and XFEL pulse employed in this experiment. **e,** 2mF_o-DF_c maps obtained from the non-pre-flashed dark-state data (gray contoured at 2σ) and mF_o-DF_c map (green (negative) and red (positive) contoured at ±5σ) of a region around the Chl_{D1} site, before locating the two water molecules in the vicinity of Chl_{D1}. **f,** Stereo image of 2mF_o-DF_c maps in a region around OEC, with the gray and cyan contoured at 2σ and 8σ, respectively.

Figure 2 | Isomorphous difference Fourier map between the 2F-state and the dark-state of the pre-flashed samples. a-b, Isomorphous difference Fourier map between the 2F-state and the dark-state is shown in green (positive) and red (negative) contoured at ±4.5σ with a top view from the luminal side **(a),** a side view perpendicular to the membrane normal **(b)**. Cycles in red dotted lines in **(b)** indicate two regions where significant difference peaks were observed. **c,** A stereo image of the region boxed by black dashed lines in **b** (labeled c). **d,** Histogram of the isomorphous difference Fourier map, showing apparent deviations of the difference Fourier peaks from the Gaussian distribution of the noise level beyond around ±5σ. Two arrows indicate the possible largest noise level.

Figure 3 | Isomorphous difference Fourier map and structural changes around the Q_B/non-heme iron-binding site and OEC. Isomorphous difference Fourier maps between the 2F-state and the dark-state of the pre-flashed samples are shown in green (positive) and red (negative) contoured at ±4.5σ for **a, b**, and ±5.0σ for **c, d**. D1 and D2 subunits are colored in yellow and cyan in the S_1 state, and green in the S_3 state. **a and b,** Structural changes around the Q_B and non-heme iron-binding site. **c and d,** Structural changes around OEC. In **b-d**, movement of residues/groups are indicated by black arrows, and the insertion or displacement of water (oxygen) are indicated by arrows with dashed lines. **e,** Structure of the Mn_4CaO_5 cluster after 2F-illumination superimposed with the $2mF_o$-DF_c map countored at 10.1σ. **f,** Position of the newly inserted oxygen atom O6 relative to its nearby atoms.

Figure 4 | Possible mechanism for O=O bond formation and the proton transfer pathways from OEC. The circle in dotted line indicates displacement of the water molecule (W665) next to the water (W567) H-bonded to O4, which results in the approaching of W567 to O4. Arrows in dotted line indicate possible pathways for proton transfer. The two pairs of oxygen atoms cycled by dashed red lines represent the possible sites for O=O bond formation, with the one between O5 and O6 have a much shorter distance and therefore representing the real site for O=O bond formation.

Methods

No statistical method was used to predetermine the sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

Preparation and crystallization of photosystem II.

Highly active PSII was isolated from *Thermosynechococcus vulcanus* and crystallized as described previously with slight modifications[4,27,28]. The final PSII core dimers were suspended in 20 mM Mes (pH 6.0), 10 mM NaCl, 3 mM $CaCl_2$, and the final crystallization buffer contained 20 mM Mes, 20 mM NaCl, 40 mM $MgSO_4$, 10 mM $CaCl_2$, 5-7% polyethylene glycol 1,450, 0.85% n-heptyl-β-D-thioglucopyranoside (Dojindo). No re-crystallization procedure was applied. We screened various sizes of PSII crystals and post-crystallization treatment conditions (see below) to prepare suitable samples for TR-SFX by XFEL, and found that too small crystals did not diffract to a high resolution, whereas larger crystals gave rise to a lower efficiency of the S*i*-state transition induced by the laser excitations. The optimal crystal size was determined to have a maximum length of 100 μm, which diffracted up to a resolution of 2.1 Å by a SACLA-XFEL pulse (Extended Data Fig. 1) and a final population of 46% S_3 state upon 2F-illumination (see below).

All of the procedures for the preparation, crystallization, pre-flash illumination and diffraction experiments were conducted in the dark or very dim green light. To prepare a large amount of micro-sized crystals, crystallization was performed in 1.5 ml micro centrifuge tubes with a sample volume of 50 μl at a concentration of 2.3 mg chlorophyll/ml at 20ºC. The crystals appeared in a few hours; when the crystal size

reached to a maximum length of 100 μm, 50 μl of the crystallization buffer in which the concentration of PEG1,450 was increased by 1-2% from the crystallization condition was added to stop further growth of the crystals (see Extended Data Fig 1a for a picture of typical crystals). The crystals were washed several times with this buffer to remove the PSII samples that were not crystallized, and stored overnight.

Prior to the XFEL experiments, the crystals were transferred to a fresh mother liquid containing 10 mM potassium ferricyanide as an electron acceptor for the pre-flashing samples, and 2 mM potassium ferricyanide for the samples without pre-flash (non-pre-flash). The concentration of "cryo-protectant" was then increased finally to 22% glycerol, 9% PEG1,450 and 9% PEG5,000 MME by a stepwise replacement of the mother solution in 1.5 hours. Finally, the crystals were carefully mixed with a grease matrix, Super Lube nuclear grade grease (Synco Chemical Co.), and loaded into a high viscosity micro-extrusion injector as described previously[13].

Although SFX can be performed at room temperature and there is no need to freeze the crystals, we found that this post-crystallization procedure to increase the "cryo-protectant" concentration was important for obtaining good diffracting PSII crystals. When the post-crystallization procedure was not adequate, some crystals gave rise to larger unit cell dimensions of a = 129.1 Å, b = 228.8 Å and c = 305.4 Å (Extended Data Fig. 6a-c). Diffraction spots from the crystals with this larger unit cell were found to be lower than 3.0 Å resolution probably due to the loose crystal packing. Indeed, we found that the PSII dimer in this crystal packing harbored two PsbY subunits (one PsbY per monomer PSII) (Extended Data Fig. 6e). On the other hand, only one PsbY was found in the PSII dimer (one monomer contained PsbY whereas the other monomer did not) in the structure with the unit cell dimensions of a = 126.5 Å, b =

231.2 Å and c = 287.5 Å analyzed in this study (Extended Data Fig. 6f). When the two structures were superimposed, one of the two PsbY subunits in the dimer structure with the larger crystal packing indeed interfered with the adjacent monomer in the structure with the smaller crystal packing (Extended Data Fig. 6d-f).

Pre-flash illumination of PSII crystals

In order to decrease the possible contamination of the S_0 state in the dark-adapted crystals, PSII micro-crystals were illuminated with one pre-flash before the post-crystallization treatment. This was carried out as follows. An aliquot of 100 µl solution containing micro-crystals of PSII was transferred into a dialysis button, and the pre-flash was provided by a Nd:YAG laser (Minilite-I, Continuum) at 532 nm with a diameter of 7 mm (large enough to cover the entire sample area) at an energy of ~52 mJ/cm^2 at the sample position. After the pre-flash illumination, the sample was stored for 1-3 hours in the dark condition while being transferred into the cryo-protectant conditions. This dark-incubation time was long enough to allow the higher S*i*-states (S_2 and S_3) to decay into the S_1 state, and therefore minimize the possible contamination of the S_0 state in the dark-adapted crystals, if there are any.

Estimation of the S_3 population in the PSII crystals by ATR-FTIR measurements

The population of the S_3 state after 2 flashes in the PSII micro-crystals was estimated using light-induced Fourier transform infrared (FTIR) difference spectroscopy combined with the attenuated total reflection (ATR) method. PSII crystals in the same buffer as the SACLA experiments (including 10 mM potassium ferricyanide) were loaded onto a three-reflection silicon prism and then sealed with a transparent plate and

a silicone rubber spacer. For the measurement of standard spectra, the PSII core solution was loaded on the silicon prism, and the sample temperature was maintained at 20˚C. FTIR spectra were recorded at 4 cm^{-1} resolution using a Bruker IFS-66/S spectrophotometer equipped with an MCT detector. Flash illumination was provided by a Q-switched Nd:YAG laser (Quanta-Ray INDI-40-10; 532 nm, ~7 ns FWHM, ~21 mJ cm^{-2}). Saturation of the laser flash was confirmed by checking the power dependence of the FTIR signal. After two pre-flashes with subsequent dark adaptation for 30 min, two flashes with an interval of 10 sec were applied to the sample and FTIR difference spectra were measured upon each flash. We confirmed that the higher S-states decay very little during this 10 sec interval. This measurement was repeated 10 times to increase the signal-to-noise ratio of the spectra, with each measurement separated by 30 min dark incubation.

The FTIR difference spectra obtained upon the 1st (a) and 2nd (b) flashes of PSII complexes in solution (black lines) and crystals (red lines) are shown in Extended Data Fig. 2. The spectra are normalized by the intensity of the amide II band, which reflects the protein amount. The efficiencies of the S-state transitions in the PSII crystals were estimated following the method previously described[29,30]. The 1st- and 2nd-flash spectra in the PSII crystal, $f_1(\nu)$ and $f_2(\nu)$, respectively, were fitted with linear combinations of the 1st- and 2nd-flash spectra in solution, $F_1(\nu)$ and $F_2(\nu)$, respectively, as standard spectra: $f_1(\nu) = c_{11}F_1(\nu)$; $f_2(\nu) = c_{21}F_1(\nu) + c_{22}F_2(\nu)$, where c_{11}, c_{21}, and c_{22} are the coefficients of linear combination. The least-squares fitting was performed in the symmetric COO^- region (Extended Data Fig. 2b). The coefficients were estimated to be c_{11} = 0.73 ± 0.02, c_{21} = 0.27 ± 0.01, and c_{22} = 0.64 ± 0.02. The efficiencies of the $S_1 \rightarrow S_2$ and $S_2 \rightarrow S_3$ transitions are expressed as $\alpha_0 c_{11}$ and $\alpha_0(c_{22}/c_{11})$, respectively, and

the population of the S_3 state after the 2nd flash is calculated as $\alpha_0^2 c_{22}$. Here, α_0 is the average efficiency of S-state transitions in solution and it was determined to be 0.85 ± 0.01 from the oscillation pattern of the intensity at 1400 cm^{-1} [31,32] obtained by 12 consecutive flashes. Based on this, the population of the S_3 state in the PSII micro-crystals after 2 flashes is estimated to be 0.46 ± 0.03.

Diffraction experiment at SACLA-XFEL

Single-shot XFEL data collection was performed using femtosecond X-ray pulses from the SACLA at BL3. The pulse parameters of SACLA were as follows: pulse duration, 2-10 fsec; X-ray energy, 7 keV; energy bandwidth, 0.5% (FWHM); pulse flux, ~7 × 10^{10} photons/pulse; beam size 3.0 μm (H) × 3.0 μm (W); repetition rate, 30 Hz.

The PSII crystals mixed with grease were loaded into an injector with a nozzle diameter of 150 μm, and set in a diffraction chamber filled with helium gas in a setup called Diverse Application Platform for Hard X-ray Diffraction in SACLA (DAPHNIS)[33]. The flow rate was set to 5.6 μl/min (5.28 mm/sec) for the 2F-data and 2.8 μl/m (2.64 mm/sec) for the dark-data. The diffraction patterns were recorded using a multiport CCD detector[34]. Because the excitation laser pulses were provided at 10 Hz and the XFEL pulses had a repetition rate of 30 Hz, each “pump-on” image was followed by two “pump-off” images, which were recorded separately. The "pump-on" images were used to analyze the 2F-state structure, whereas the diffraction data for the dark-state was collected by a separate run.

To advance the PSII samples to the S_3-state, two consecutive excitation laser flashes were provided from two separate Nd:YAG laser sources (Minilite-I, Continuum) to the sample at t=0 msec and t=10 msec (i.e., the two flashes were separated by 10 msec) for

the pre-flashed samples, and t=0 msec, t=5 msec (separated by 5 msec) for the non-pre-flashed samples, respectively. To ensure sufficient excitation, each of the pump lasers were split into two beams, and each beam from the same pump laser (pump 1 or pump 2) was focused on the same sample point with an angle of 160^{o} with respect to each other (a nearly counter-propagating geometry). The two beams from pump 1 were used for the first flash illumination, whereas the other two beams from pump 2 were used for the second flash illumination. The pump focal diameter of all beams was set to 240 μm (top-hat) and its energy was 42 mJ/cm^2 from each direction. The XFEL pulses were provided 93 μm downstream from the pump focal center at a time 10 msec for the pre-flashed samples and 15 msec for the non-pre-flashed samples, respectively, following the second excitation laser (the total time following the first flash is therefore 20 msec in both cases) (Fig. 1d-e). At the flow rate of 5.6 μl/min, each pump-illuminated crystal for the "pump-on" XFEL images was separated by 528 μm, which was long enough to avoid the influence from the previous excitation lasers, as the illumination point by the last laser extends to only 120 μm from its center of illumination.

Data processing

The background of the detector was estimated by averaging dark images and subtracted from diffraction patterns. Diffraction images were filtered by the program *Cheetah*[35,36] adapted to SACLA, and processed by the program *CrystFEL*[37]. The parameters "min-snr", "thresholds", "min-gradient" used for peak-detection during spot-finding were as follows: 6, 500, 10,000 for the pre-flashed dark- and pre-flashed 2F-data; 5, 500 and 10,000 for the non-pre-flashed dark-data; 5, 500 and 5,000 for the non-pre-flashed

2F-data. Indexing was performed using DirAx[38] and Mosflm[39] with peak integration parameters “int-radius= 3, 5, 7”, where the unit cell information was provided to avoid integration of the diffraction images from the crystals with a longer *c* axis. Before Monte-Carlo integration, the indexed images with either a “diffraction_resolution_limit” lower than 4.2 Å or a “num_peaks” less than 400 were discarded. The numbers of total images collected, the hit images filtered by *Cheetah*, the indexed images and the number of images used for refinement for pre-flashed samples were as follows: 408,071, 76,047, 64,985, 27,497 for the dark data; 273,550, 60,885, 51,482, 21,680 for the 2F-data. The corresponding numbers for the non-pre-flashed samples are 462,343, 70,083, 54,956 and 22,341 for the dark data; 876,874, 165,463, 63,711 and 23,903 for the 2F-data (Extended Data Table 1). The Lorentz factor for still snapshots was applied manually to the averaged intensities of the pre-flashed dark and 2F datasets[40]. The statistics for the data collection were given in Extended Data Tables 1 and 2, which showed that our data for both dark and 2F-states had a resolution of 2.35 Å for the pre-flashed samples, and 2.5 Å for the non-pre-flashed samples, based on the cut-off value of around 50% for $CC_{1/2}$. We should point out that the overall multiplicities of all our data are very high, and even at the highest resolution shell, the multiplicity well exceeded 500. Together with the facts that: (i) the values of $CC_{1/2}$ in the highest resolution shells are reasonably high; (ii) $CC_{1/2}$ decreases gradually without any abnormal disrupt from the low resolution to higher resolution shells; and (iii) the value of $CC_{1/2}$ in the highest resolution shell is reasonably consistent with the value of I/σ (I) (theoretically 0.5 of $CC_{1/2}$ corresponds to 2.0 of I/σ (I), and any substantial deviations from these values are indicators of systematic bias and/or problems in the error model, see ref. 41), we consider that the quality of the data in the present study is reasonably

high to allow us to reveal the small structural changes induced by the flash illuminations.

Structural refinement for the pre-flashed dark-adapted S_1 state, and non-pre-flashed S_1 state and 2F state

The initial phases up to 4 Å resolution were obtained by molecular replacement with the program *Phaser* in the CCP4 suite[42] using the 1.95 Å resolution XFEL structure of PSII (PDB accession code 4UB6[5]) as the search model, in which OEC and its direct ligands, Q_B, waters and glycerol molecules were omitted. After a few cycles of the rigid body refinement and subsequent real space density modification using solvent flattening, histogram matching and non-crystallographic symmetry averaging with the program DM in the CCP4 suite[42], the electron density map obtained showed features clear enough to allow us to build the model with confidence. Structural refinement was performed with the program *Phenix*[43] and the model was manually modified with the program *COOT*[44]. After a few cycles of the restrained refinement, water molecules were placed in positions corresponding to the water molecules in the higher-resolution structure where positive peaks higher than 3.5σ in the mF_o-DF_c map were clearly identified. Then, based on the resulted mF_o-DF_c map, water molecules and glycerol molecules were additionally located when the positive peaks higher than 3σ were found. Q_B, OEC and its direct ligand residues were modeled in the final step. In our previous work of XFEL structural analysis at the 1.95 Å resolution[5], the exact positions of the oxo-bridges were identified in the mF_o-DF_c map by omitting the oxo-bridged oxygen atoms; however, when the restrained refinement was performed in the same way in this study, the temperature factors of the manganese and/or calcium atoms in OEC were not

converged, and gave rise to very high values due to the limited resolution of 2.35 - 2.5 Å. Thus, we built the OEC structure using the geometric restraints based on the Mn_4CaO_5 cluster in the 1.95 Å resolution XFEL structure, and applied tight distance restraints of σ = 0.02 Å to the Mn-O bonds, Ca-O bonds and Mn-ligand residue distances during the refinement. No restraints were given to the Mn-Mn distances and Mn-Ca distances. The *R*-factor and *R*-free values obtained were 0.133, 0.171 for the pre-flashed dark-state, and their corresponding values were 0.139, 0.186 for the dark-state, 0.139, 0.187 for the 2F-state of the non-pre-flashed samples (Extended Data Table 1).

Structural refinement for the pre-flashed 2F state

Given the population of the S_3 state estimated from the FTIR measurement, the diffraction data obtained from the 2F-illuminated sample is expected to contain those partly from the S_2 and S_1 states of PSII, which means that the resulted electron density would be a mixture of these S-states including the S_3 state. Even with a high resolution dataset, the structural refinement of the mixed states (or mixed structures) would be challenging, especially when the mixed structures are very similar as in the case of PSII. As we did not have the structure corresponding to the S_2 state, we refined the pre-flashed 2F dataset as a mixture of the S_1 state structure and 2F state structure[45]. The occupancies for the S_1 structure and 2F state structure was set to 0.2 and 0.8, respectively, on the basis of the distributions of the temperature factors. The region in which the structure was refined in multiple states were selected based on the large peaks observed in the isomorphours difference Fourier map, and the rest was refined as a single conformation. As a result, thirty-five amino acid residues, five water molecules,

one non-heme iron, one Q_B and one OEC per monomeric PSII were modeled as multiple states. During the restrained refinement, the coordinate for the S_1 state structure was fixed. Relatively loose geometric restraints for distance, angle and plane with relatively high sigma values (three to ten times larger than the default values) were applied to OEC and Q_B during the refinement. A large positive peak of 10.3σ for monomer-A and 9.5σ for monomer-B found at a position with distances of 1.5 Å and 2.3 Å to O5 and Mn1D, was modeled as a new oxo-oxygen O6 with an occupancy of 0.4, which gave rise to a temperature factor similar with its nearby atoms. This is also consistent with the population of the S_3-state estimated from the FTIR results. The *R*-factor and *R*-free values thus obtained were 0.129, 0.176, respectively.

Data availability

The structure factors and atomic coordinates have been deposited in Protein Data Bank (PDB) with accession numbers 5WS5 and 5WS6 for the pre-flashed dark-stable S_1 state and 2F states; 5GTH and 5GTI for the non-pre-flashed dark-stable S_1 state and 2F-state, respectively. All other data associated with this manuscript are available from the authors on reasonable request.

27. Shen J.-R. & Inoue Y. Binding and functional properties of two new extrinsic components, cytochrome *c*-550 and a 12 kDa protein, in cyanobacterial photosystem II. *Biochemistry* **32,** 1825-1832 (1993).
28. Shen, J.-R. & Kamiya, N. Crystallization and the crystal properties of the oxygen-evolving photosystem II from *Synechococcus vulcanus*. *Biochemistry* **39**, 14739–14744 (2000).
29. Noguchi, T. & Sugiura, M. Flash-induced FTIR difference spectra of the water oxidizing complex in moderately hydrated photosystem II core films: effect of

hydration extent on S-state transitions. *Biochemistry* **41,** 2322-2300 (2002).
30. Suzuki, H., Sugiura, M., & Noguchi, T. pH dependence of the flash-induced S-state transitions in the oxygen-evolving center of photosystem II from *Thermosynechoccocus elongatus* as revealed by Fourier transform infrared spectroscopy. *Biochemistry* **44,** 1708-1718 (2005).
31. Noguchi, T. & Sugiura, M. Flash-induced Fourier transform infrared detection of the structural changes during the S-state cycle of the oxygen-evolving complex in photosystem II. *Biochemistry* **40,** 1497-1502 (2001).
32. Suzuki, H., Sugiura M. & Noguchi T. Monitoring proton release during photosynthetic water oxidation in photosystem II by means of isotope-edited infrared spectroscopy. *J. Am. Chem. Soc.* **131,** 7849-7857 (2009).
33. Tono, K. *et al.* Diverse application platform for hard X-ray diffraction in SACLA (DAPHNIS): application to serial protein crystallography using an X-ray free-electron laser. *J. Synch. Rad.* **22**, 532–537 (2015).
34. Kameshima, T. *et al.* Development of an X-ray pixel detector with multi-port charge-coupled device for X-ray free-electron laser experiments. *Rev. Sci. Instrum.* **85,** 033110. doi: 10.1063/1.4867668. (2014).
35. Barty, A. *et al.* Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Cryst.* **47**, 1118–1131 (2014).
36. Nakane, T. *et al.* Data processing pipeline for serial femtosecond crystallography at SACLA. *J. Appl. Cryst.* **49**, 1035–1041 (2016).
37. White, T. A. *et al.* CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Cryst.* **45**, 335–341 (2012).
38. Duisenberg, A.J.M. Indexing in single-crystal diffractometry with an obstinate list of reflections. *J. Appl. Cryst.* **25,** 92–96 (1992).
39. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. IMosflm: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Cryst.* **D67,** 271–281 (2011).
40. Kabsch, W. Processing of X-ray snapshots from crystals in random orientations. *Acta Cryst.* **D70,** 2204-2216 (2014).
41. Karplus, P. A. & Diederichs, K. Assessing and maximizing data quality in macromolecular crystallography. *Curr. Opin. Struct. Biol.* **34,** 60-68 (2015).
42. Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Cryst.* **D50**, 760–763 (1994).
43. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* **D66**, 213–221 (2010).

44. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst.* **D60**, 2126–2132 (2004).
45. Nango, E. *et al.* A three dimensional movie of structural changes in bacteriorhodopsin. *Science* **354,** 1552-1557 (2016).

Main Figures

Fig. 1

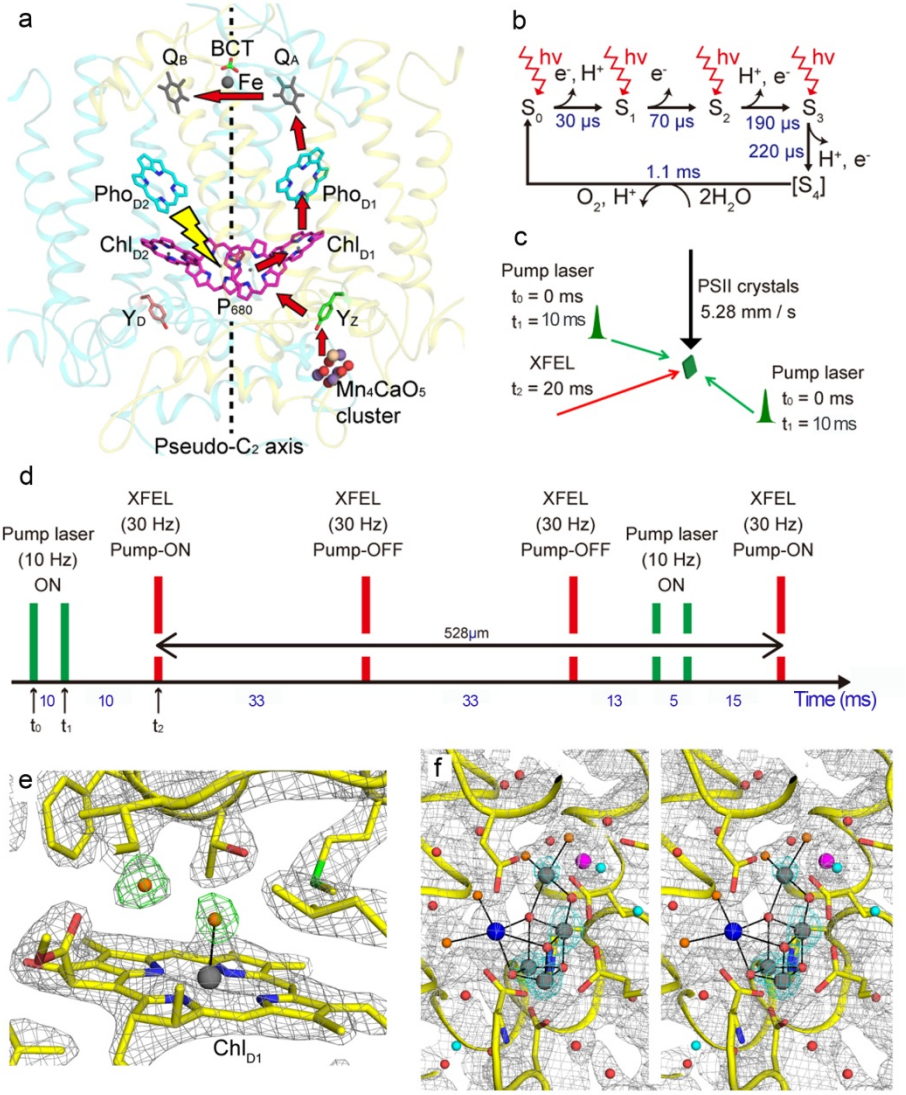


Fig. 2

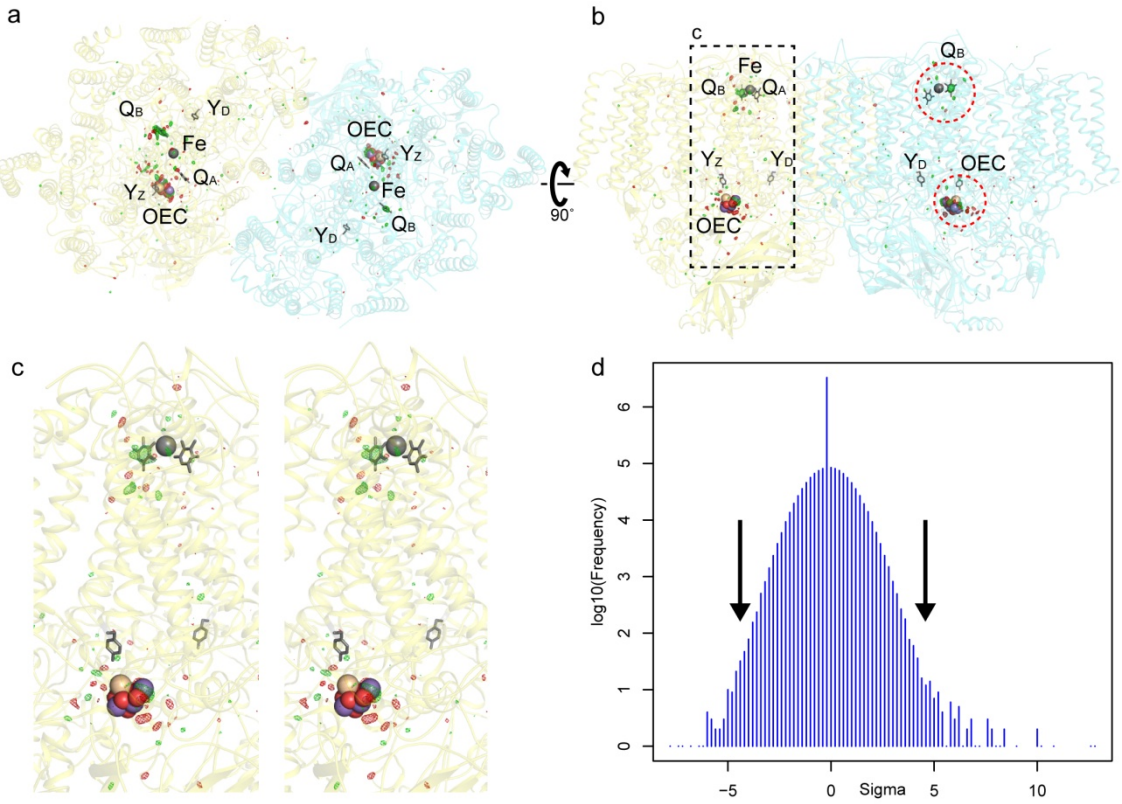


Fig. 3

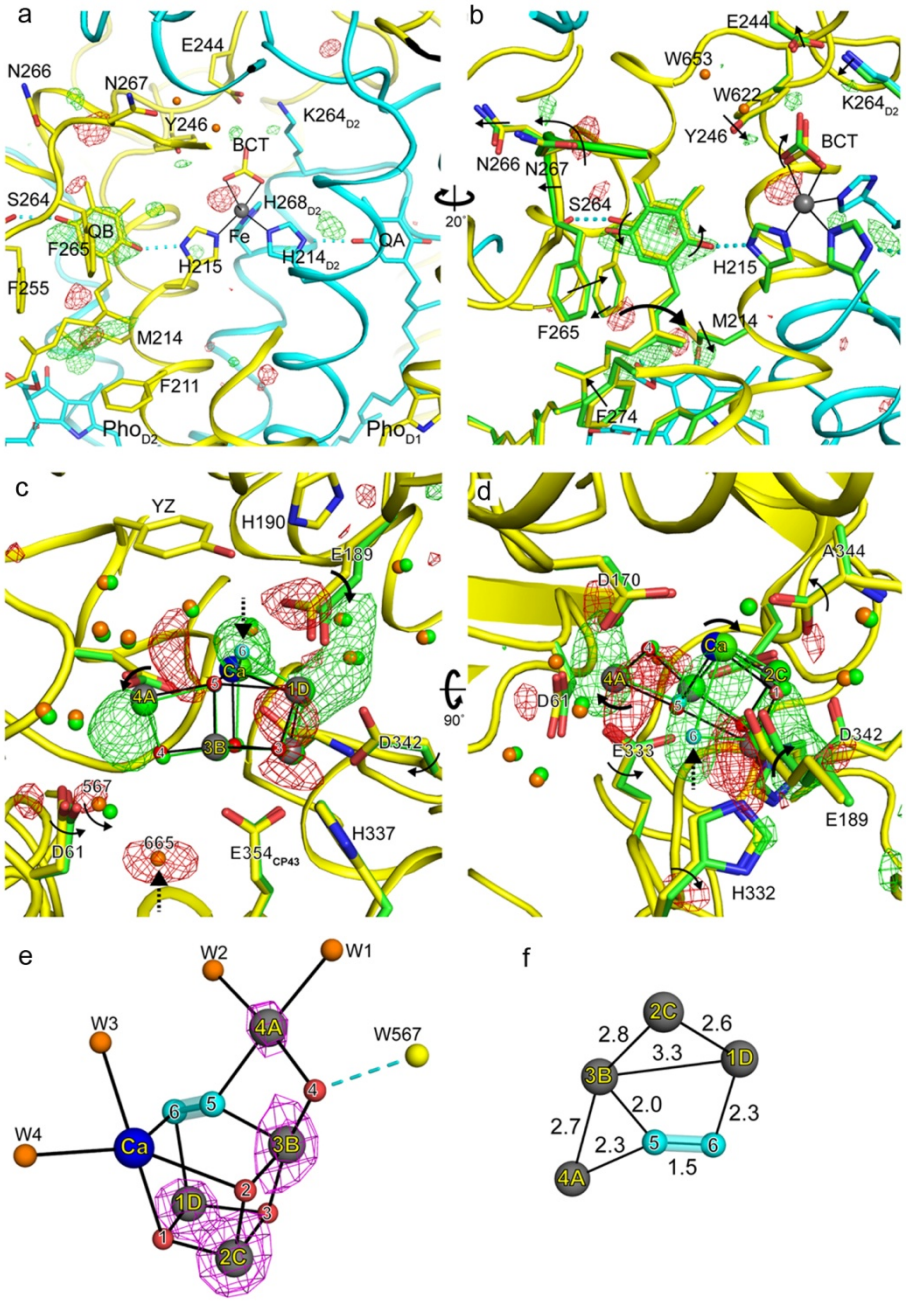
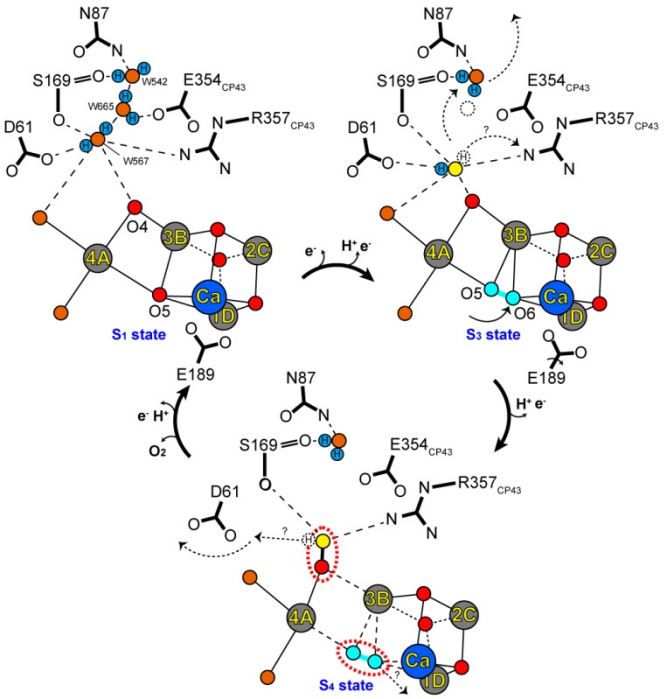

a
E244
N266
N267
Y246
K264D2
BCT
S264
H268D2
QB
Fe
QA
F265
H215
H214D2
F255
M214
F211
PhoD2
PhoD1
b
E244
W653
W622
K264D2
Y246
BCT
N266
N267
S264
H215
F265
M214
F274
20°
c
YZ
H190
E189
Ca
4A
1D
3B
D342
567
D61
665
E354CP43
H337
d
A344
D170
Ca
4A
2C
D61
E333
D342
E189
H332
90°
e
W2
W1
W3
4A
W567
W4
Ca
3B
1D
2C
f
2C
2.6
2.8
3.3
1D
3B
2.0
2.3
2.7
2.3
5
6
1.5
4A

Fig. 4


N87
S169
W542
E354CP43
W665
D61
R357CP43
W567
O4
3B
4A
2C
Ca
O5
1D
S1 state
E189
e- H+
O2
e-
H+ e-
O6
S3 state
S4 state

Legends for Extended Data Figures and Tables

Extended Data Figure 1 | Micro-sized crystals of PSII, its diffraction image and quality of the diffraction datasets. **a**, Micro-sized crystals of PSII used for TR-SFX. **b**, A diffraction image from a micro-sized PSII crystal obtained with a single pulse from SACLA-XFEL. **c-f,** Quality of the datasets. CC1/2 **(c)**, R_{split} **(d)**, multiplicity **(e)** and $\langle I/\sigma(I)\rangle$ **(f)** vs resolution for the dark-data and the 2F-data of the pre-flashed samples.

Extended Data Figure 2 | Measurement of the S_3 state population upon 2F-illumination by FTIR. a, ATR-FTIR difference spectra upon the 1st (a) and 2nd (b) flashes of the PSII crystal (red lines) and PSII solution (black lines). **b**, Result of the least-squares fitting analysis of the 2nd-flash spectrum of the PSII crystal in the symmetric COO^- region (1470-1300 cm^{-1}). The spectrum of the PSII crystal (red line) was fit with a linear combination of the 1st- and 2nd-flash spectra of the solution sample, $F_1(\nu)$ and $F_2(\nu)$, respectively (black lines in panel **a**). The resultant fitting spectrum (gray line) and the two components, $0.27F_1(\nu)$ (blue line) and $0.64F_2(\nu)$ (green line), are also shown in panel **b**.

Extended Data Figure 3 | Anomalous signals obtained with the XFEL beam. a-b, Anomalous difference Fourier maps for OEC atoms **(a)**, a region around OEC **(b)**, and a newly identified Ca binding site in CP43 **(c)**, of the non-pre-flashed dark-data are shown. The maps are shown in cyan and magenta contoured at 4σ and 12σ **(a)**, gray contoured at 3σ **(b)** and cyan contoured at 4σ **(c)**. **d**, Histogram analysis of the anomalous map of the non-pre-flashed dark data.

Extended Data Figure 4 | Electron density maps and structural changes around Y_D between the non-pre-flashed 2F and dark-datasets. a, Isomorphous difference Fourier map between the non-pre-flashed 2F-state and dark-state datasets shown in green (positive) and red (negative) contoured at ±5σ, superimposed with the non-pre-flashed dark-structure and 2F-structure. Non-pre-flashed dark and 2F structures are colored in yellow and green, respectively. The two different positions of W508 observed in the 1.95 Å resolution XFEL structure (W508I and W508II) are shown as

black dots. **b-c,** 2mF_o-DF_c map and mF_o-DF_c map of the non-pre-flashed dark-state **(b)** and 2F-state **(c)**. 2mF_o-DF_c map is colored in gray contoured at 1.5σ, and mF_o-F_c map is colored in cyan (positive) and brown (negative) contoured at ±4σ.

Extended Data Figure 5 | Electron density maps for the O4-water chain. The mF_o-DF_c maps calculated for the O4-water chain by omitting the four water molecules W567, W665, W542 and W546 are shown for the non-pre-flashed dark-state (**a** and **b**) and 2F-state (**c** and **d**). The maps shown are for monomer-A (**a** and **c**) and monomer-B (**b** and **d**) in cyan (positive) and orange (negative) contoured at ±3σ.

Extended Data Figure 6 | Distribution of the unit cell parameters and crystal packing of PSII for the non-pre-flashed samples. a-c, Distributions of unit cell parameters for the dark-state images **(a)**, and 2F-state images **(b)** from the non-pre-flashed samples and the dark-images when the post-crystallization procedure was not adequate **(c)** are shown with their average lengths and the standard deviations. No large differences were observed between the dark-images and the 2F-images, whereas a longer *c* axis was observed in **(c)**. **d-f,** Comparison of the crystal packing between the dark-state and the dataset with a longer *c* axis is shown in a view direction perpendicular to the *bc* plane. Color codes: PSII and the crystallographic symmetric molecules in the dark state, cyan and blue; PSII, PsbY subunits in PSII and the crystallographic symmetric molecules in the dataset with a long *c* axis, gray, red and khaki; Note that PsbY has considerable steric hindrance with the adjacent PSII molecule in the crystal packing of the dark data **(e)**.

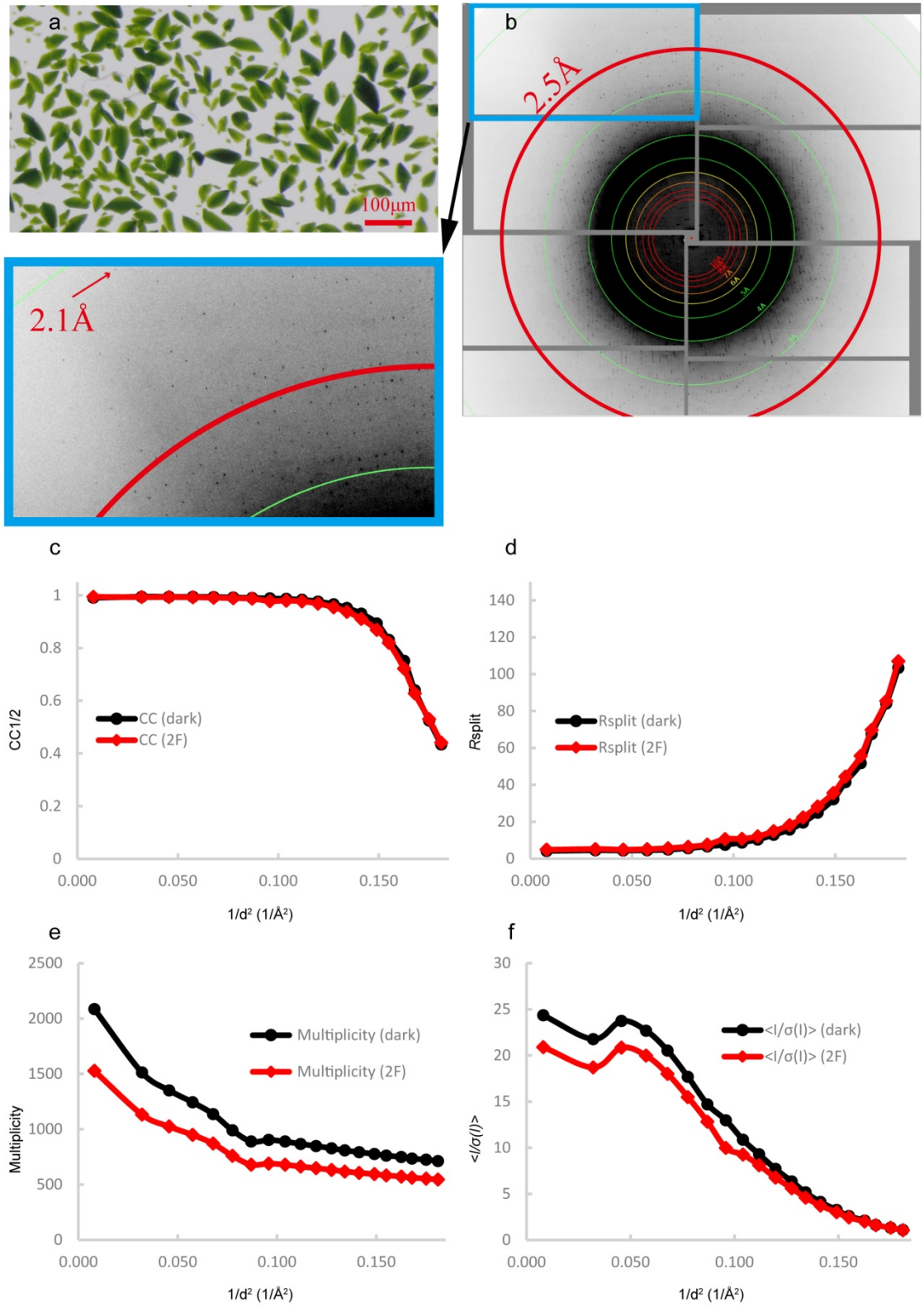
Extended Data Table 1 | Statistics of the diffraction data and structural refinements for the dark-state and 2F-state datasets of both pre-flashed and non-pre-flashed samples, collected at SACLA (a), and the R*iso* values between the different data sets (b).

Extended Data Table 2 | Statistics for the processing of the diffraction data for the dark-state (a), 2F-state (b), pre-flashed dark-state (c), and pre-flashed 2F-state collected at SACLA.

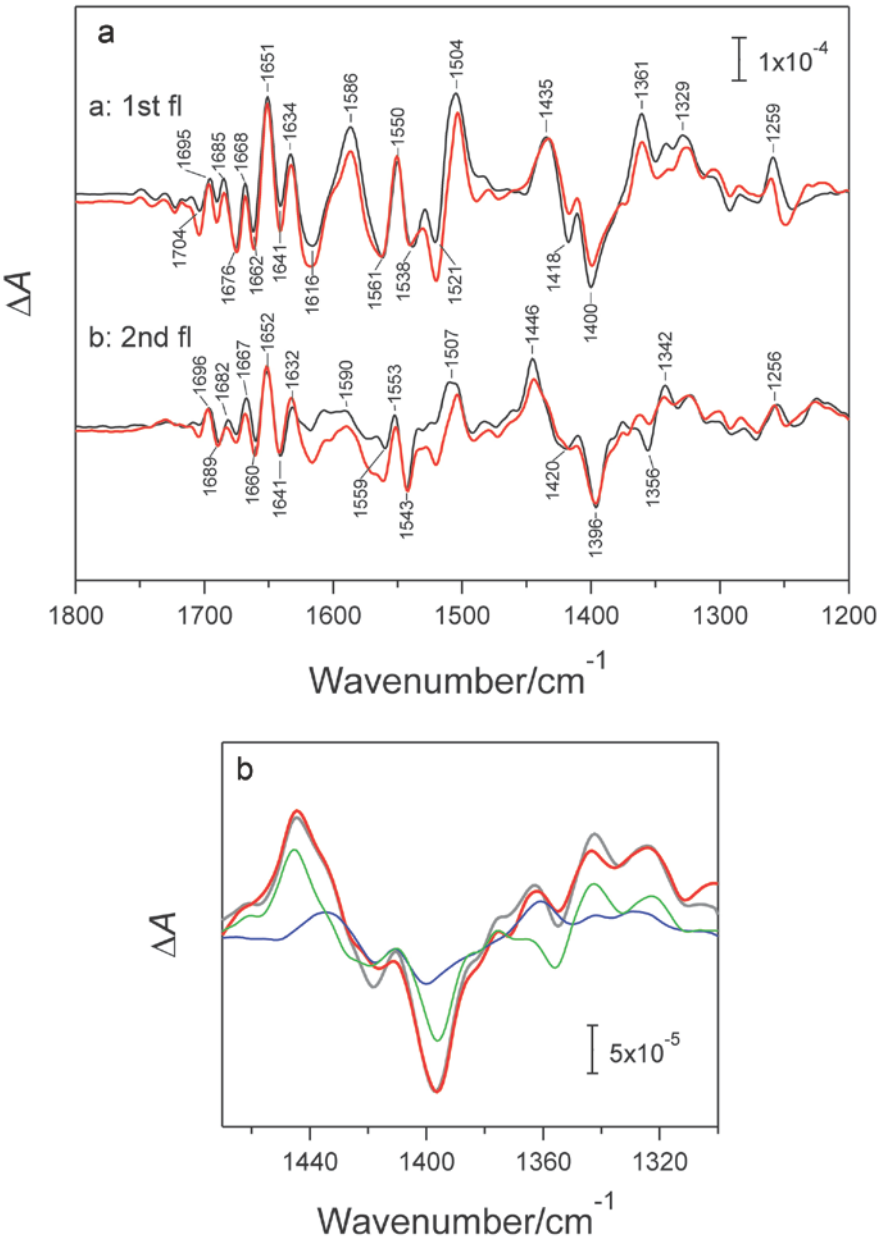
Extended Data Table 3 | List of the top 30 anomalous peaks identified in the non-pre-flashed dark state and 2F state datasets.

Extended Data

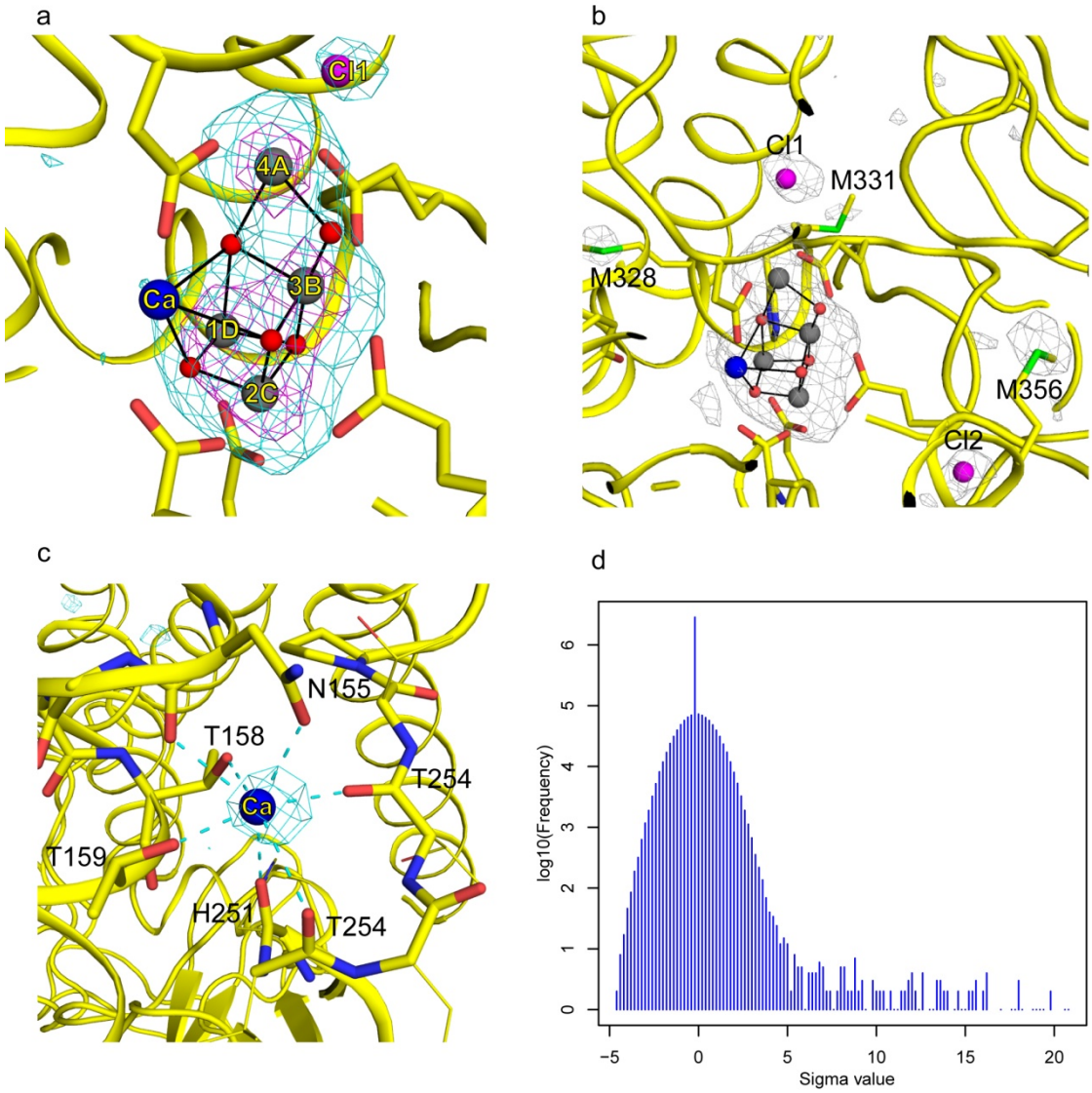
ED-Fig. 1



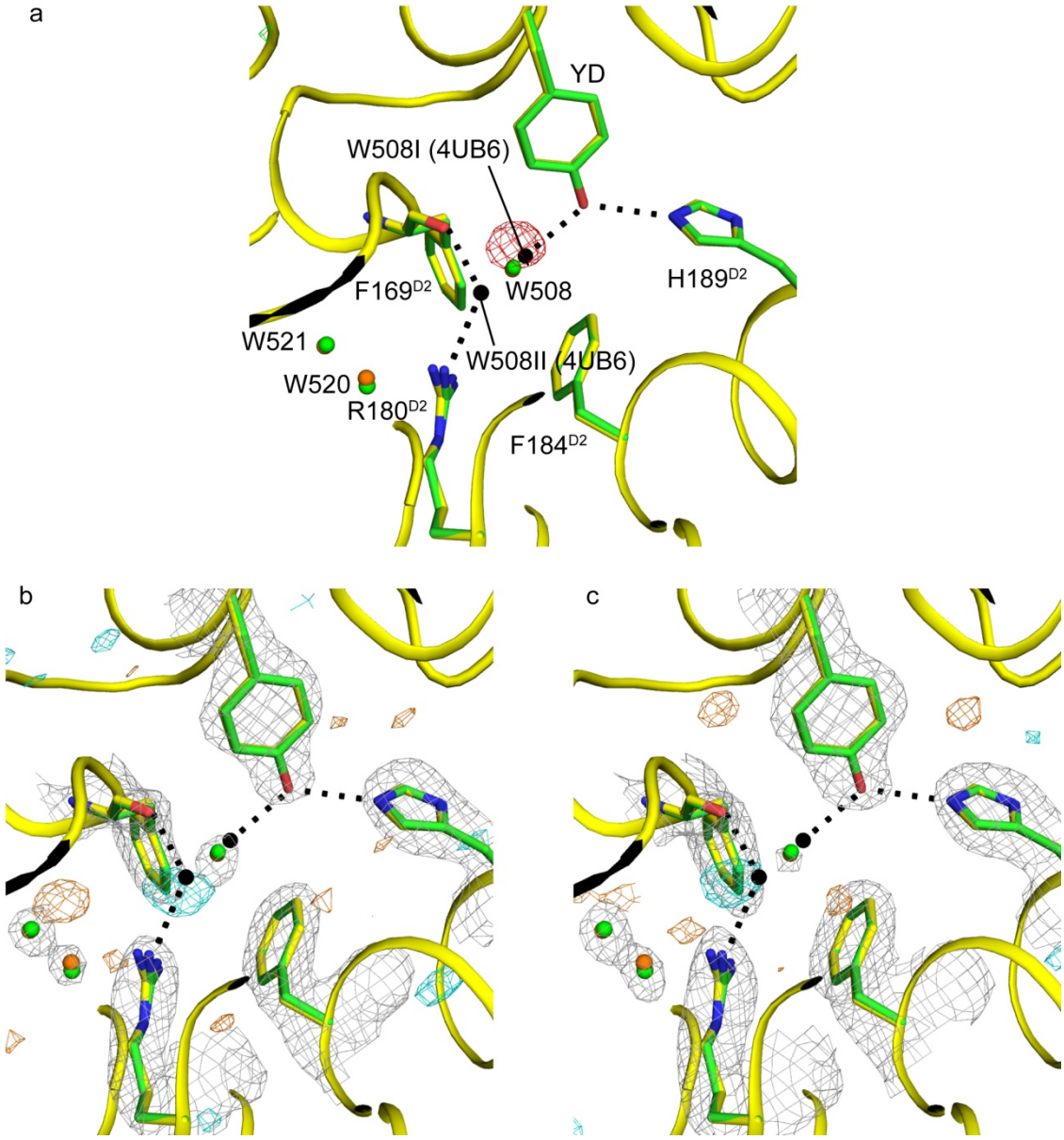
ED-Fig. 2


a
a: 1st fl
1x10-4
1695
1685
1668
1651
1634
1586
1550
1504
1435
1361
1329
1259
1704
1676
1662
1641
1616
1561
1538
1521
1418
1400
b: 2nd fl
1696
1682
1667
1652
1632
1590
1553
1507
1446
1342
1256
1689
1660
1641
1559
1543
1420
1396
1356
ΔA
1800
1700
1600
1500
1400
1300
1200
Wavenumber/cm-1
b
ΔA
5x10-5
1440
1400
1360
1320
Wavenumber/cm-1

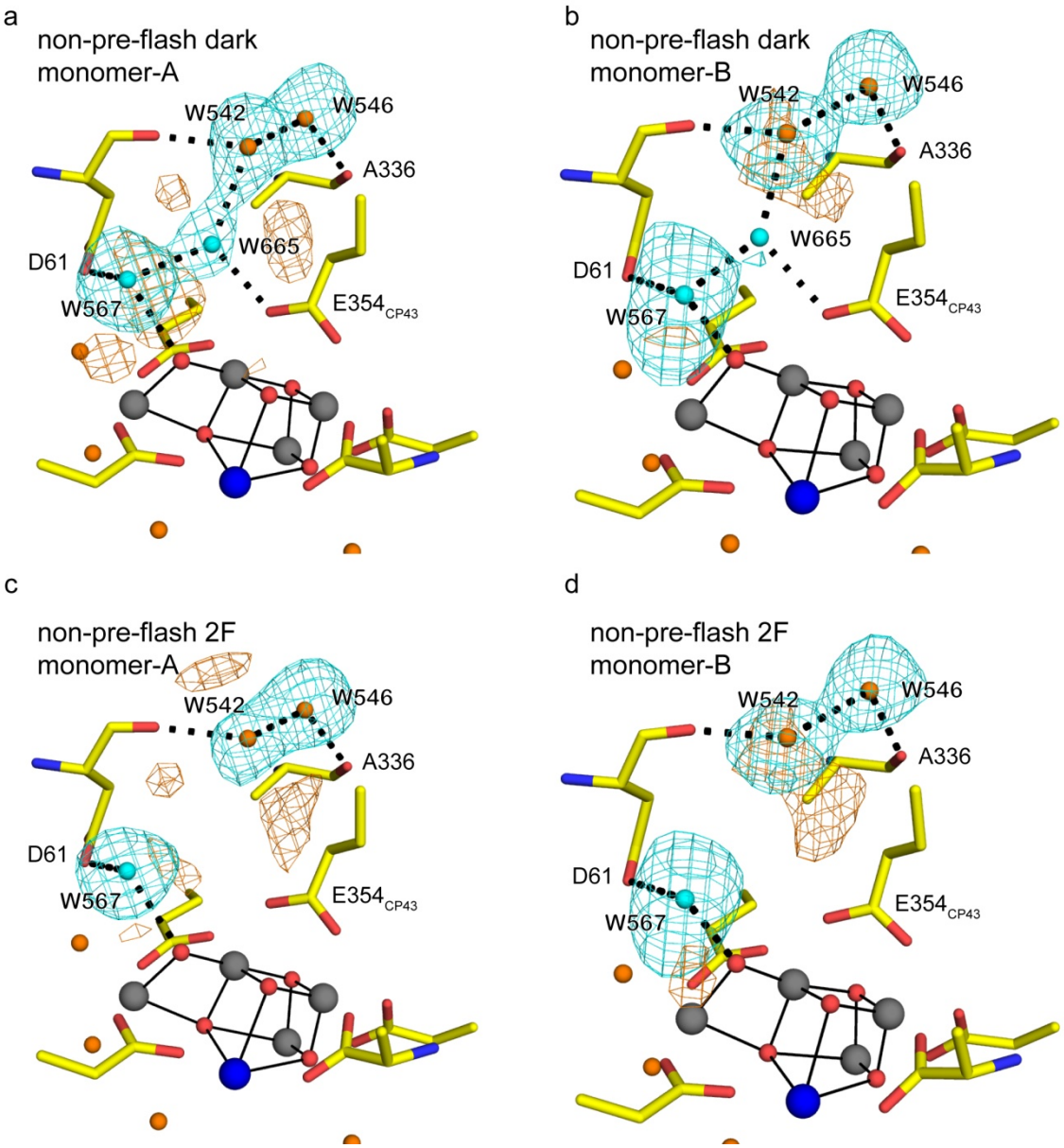
ED-Fig. 3


a
Cl1
4A
3B
Ca
1D
2C
b
Cl1
M331
M328
M356
Cl2
c
N155
T158
T254
Ca
T159
H251
T254
d
log10(Frequency)
0
1
2
3
4
5
6
−5
0
5
10
15
20
Sigma value

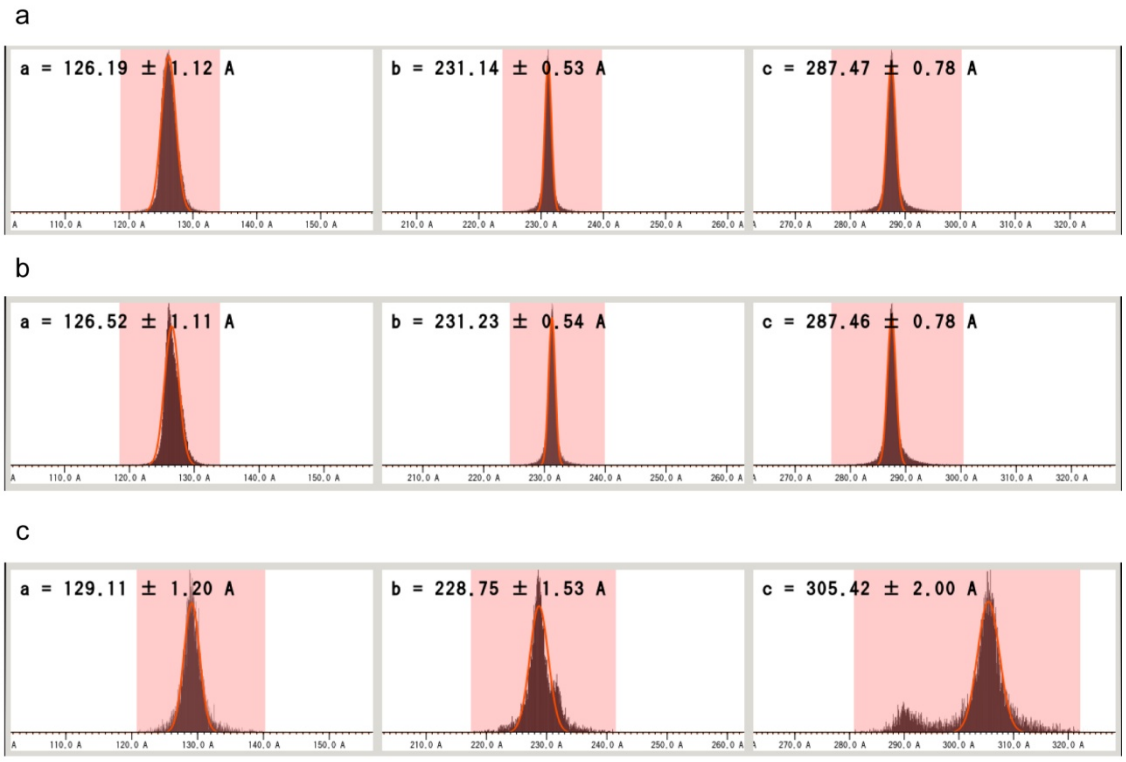
ED-Fig. 4

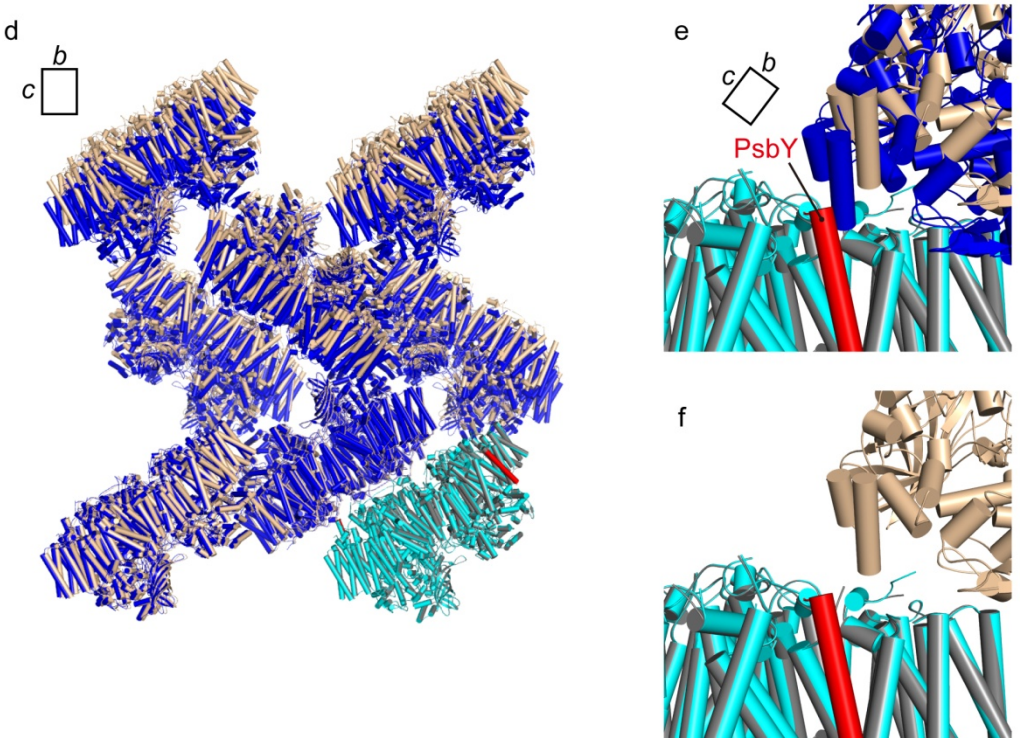

a
YD
W508I (4UB6)
H189D2
F169D2
W508
W521
W508II (4UB6)
W520
R180D2
F184D2
b
c

ED-Fig. 5


a
non-pre-flash dark
monomer-A
W542
W546
A336
D61
W665
W567
E354CP43
b
non-pre-flash dark
monomer-B
W542
W546
A336
W665
D61
W567
E354CP43
c
non-pre-flash 2F
monomer-A
W542
W546
A336
D61
W567
E354CP43
d
non-pre-flash 2F
monomer-B
W542
W546
A336
D61
W567
E354CP43

ED-Fig. 6


a
a = 126.19 ± 1.12 A
b = 231.14 ± 0.53 A
c = 287.47 ± 0.78 A
b
a = 126.52 ± 1.11 A
b = 231.23 ± 0.54 A
c = 287.46 ± 0.78 A
c
a = 129.11 ± 1.20 A
b = 228.75 ± 1.53 A
c = 305.42 ± 2.00 A


d
b
c
e
c
b
PsbY
f

ED-Table 1

a

	preflash-dark-data (5WS5)	preflash-2F-data (5WS6)	dark-data (5GTH)	2F-data (5GTI)
Data collection statistics				
Wavelength / Å	1.77	1.77	1.77	1.77
No. of collected images	408,071	273,550	462,343	876,874
No. of hit images	76,047	60,885	70,083	165,463
No. of indexed images	64,985	51,482	54,956	63,711
No. of images used for refinement	27,497	21,680	22,341	23,903
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell / Å	a=124.7, b=229.9, c=285.5	a=125.0, b=230.2, c=286.0	a=126.2, b=231.1, c=287.5	a=126.5, b=231.2, c=287.5
Resolution / Å	50.0 - 2.35 (2.44 - 2.35)*	50.0 - 2.35 (2.44 - 2.35)*	50.0 - 2.50 (2.60 - 2.50)*	50.0 - 2.50 (2.60 - 2.50)*
No. of unique reflections	357,086 (35,393)*	357,084 (35,393)*	290,634 (28,847)*	290,634 (28,847)*
Completeness / %	100 (100)*	100 (100)*	100 (100)*	100 (100)*
Multiplicity	978 (714)*	742 (546)*	760 (557)*	824 (605)*
R split‡	0.054 (1.033)*	0.062 (1.045)*	0.057 (0.959)*	0.055 (0.931)*
CC1/2§	0.994 (0.438)*	0.995 (0.445)*	0.997 (0.505)*	0.997 (0.535)*
mean I/σ (I)	10.5 (1.1)*	9.2 (1.1)*	11.1 (1.2)*	10.6 (1.2)*
Refinement statistics				
Resolution / Å	50.0 - 2.35	50.0 - 2.35	50.0 - 2.50	50.0 - 2.50
R factor	0.133	0.129	0.139	0.139
R free	0.171	0.176	0.186	0.187
Wilson B / Å^2	59.9	58.3	71.5	71.4
Average B (overall) / Å^2	71.9	67.2	48.2	51.6
Average B (protein) / Å^2	71.9	67.2	48.1	51.4
Average B (OEC) / Å^2	60.1	53.1	35.5	36.4
Average B (solvent) / Å^2	72.4	66.7	49.4	54.9
RMSD bond length / Å	0.008	0.009	0.008	0.008
RMSD bond angle / deg.	1.41	1.37	1.39	1.37
Ramachandran plot‖				
Favored / %	98.0	97.8	97.5	97.8
Allowed / %	1.9	2.0	2.3	2.0
Outliers / %	0.2	0.2	0.2	0.2

*Values in parenthesis indicate those for the highest resolution shells.

‡$R_{split} = \sqrt{2} \sum_{hkl} |I_{even} - I_{odd}| / \sum_{hkl} (I_{even} + I_{odd})$

§CC1/2 is the correlation coefficient between two half data sets.

‖Calculated with MolProbility.

b

	preflash-2F-data (5WS6)	dark-data (5GTH)	2F-data (5GTI)
preflash-dark-data (5WS5)	0.056 (0.209)*	0.106 (0.281)*	0.127 (0.344)*
preflash-2F-data (5WS6)	- -	0.118 (0.319)*	0.112 (0.304)*
dark-data (5GTH)		- -	0.070 (0.220)*

*Values in parenthesis indicate those for the highest resolution shells.

ED-Table 2

a

resolution (Å)	unique reflections	completeness (%)	measured reflections	redundancy	I/σ(I)	$CC_{1/2}$	R_{split} (%)
9.00	36,825	100.0	66,021,239	1,793	23.07	0.992	4.34
4.40	35,963	100.0	46,364,827	1,289	23.14	0.994	4.52
3.68	35,769	100.0	37,386,049	1,045	18.71	0.994	5.28
3.28	35,670	100.0	31,907,310	895	13.44	0.990	7.16
3.02	35,557	100.0	31,040,638	873	9.71	0.985	9.82
2.82	35,568	100.0	29,573,853	832	6.61	0.971	14.96
2.67	35,456	100.0	28,189,990	795	4.30	0.939	23.58
2.54	35,470	100.0	27,079,798	764	2.69	0.847	39.60
2.44	35,415	100.0	26,090,530	737	1.68	0.664	65.62
2.35	35,393	100.0	25,255,891	714	1.08	0.438	103.25
overall	357,086	100.0	348,910,125	977	10.51	0.994	5.40

b

resolution (Å)	unique reflections	completeness (%)	measured reflections	redundancy	I/σ(I)	$CC_{1/2}$	R_{split} (%)
9.00	36,823	100.0	48,804,995	1,325	19.82	0.994	5.00
4.40	35,964	100.0	35,278,411	981	20.35	0.993	5.05
3.68	35,769	100.0	28,642,604	801	16.40	0.991	6.10
3.28	35,670	100.0	24,378,291	683	10.99	0.984	9.19
3.02	35,556	100.0	23,732,000	668	8.45	0.979	11.55
2.82	35,568	100.0	22,611,764	636	5.83	0.960	17.34
2.67	35,456	100.0	21,561,321	608	3.89	0.920	26.85
2.54	35,470	100.0	20,719,126	584	2.52	0.826	43.08
2.44	35,415	100.0	19,940,363	563	1.63	0.642	68.97
2.35	35,393	100.0	19,310,841	546	1.09	0.445	104.47
overall	357,084	100.0	264,979,716	742	9.16	0.995	6.24

c

resolution (Å)	unique reflections	completeness (%)	measured reflections	redundancy	I/σ(I)	$CC_{1/2}$	R_{split} (%)
9.49	30,009	100.0	40,691,907	1,356	27.00	0.996	3.79
4.68	29,313	100.0	29,392,721	1,003	25.00	0.997	3.89
3.92	29,112	100.0	25,034,725	860	19.74	0.996	4.87
3.49	29,028	100.0	20,173,931	695	12.83	0.992	7.66
3.21	28,988	100.0	19,347,973	667	8.28	0.981	12.50
3.00	28,840	100.0	18,611,348	645	5.43	0.953	20.11
2.84	28,899	100.0	17,898,661	619	3.69	0.907	30.31
2.71	28,821	100.0	17,176,315	596	2.51	0.820	45.06
2.59	28,777	100.0	16,572,575	576	1.71	0.694	66.15
2.50	28,847	100.0	16,077,916	557	1.21	0.505	95.93
overall	290,634	100.0	220,978,072	760	11.09	0.997	5.67

d

resolution (Å)	unique reflections	completeness (%)	measured reflections	redundancy	I/σ(I)	$CC_{1/2}$	R_{split} (%)
9.49	30,009	100.0	43,880,782	1,462	26.79	0.996	3.58
4.68	29,313	100.0	31,866,344	1,087	24.43	0.997	3.81
3.92	29,112	100.0	27,192,707	934	19.12	0.996	4.90
3.49	29,028	100.0	21,909,015	755	12.36	0.992	7.76
3.21	28,988	100.0	21,010,918	725	7.99	0.981	12.56
3.00	28,840	100.0	20,214,259	701	5.26	0.954	20.13
2.84	28,899	100.0	19,430,046	672	3.60	0.907	30.29
2.71	28,821	100.0	18,647,894	647	2.46	0.827	44.89
2.59	28,777	100.0	17,994,968	625	1.70	0.703	65.79
2.50	28,847	100.0	17,445,736	605	1.21	0.535	93.11
overall	290,634	100.0	239,592,669	824	10.58	0.997	5.54

ED-Table 3

	dark			2F		
	peak hight (σ)	location	Monomer	peak hight (σ)	location	Monomer
1	22.7	Mn1D	B	22.2	Mn1D	B
2	20.7	Mn1D	A	21.2	Mn2C	A
3	19.0	Mn3B	A	19.7	Mn1D	A
4	18.0	Mn3B	B	19.0	Mn3B	A
5	17.7	Mn4A	A	17.1	Mn3B	B
6	17.4	Mn4A	B	16.4	Mn4A	A
7	6.6	Ca1-OEC	A	15.0	Mn4A	B
8	6.3	CL1	B	6.3	CL1	A
9	6.3	Ca-PsbO	B	5.9	Cys44-PsbO	A
10	6.2	Cys44-PsbO	A	5.9	Ca1-CP43	B
11	6.2	Met396-CP43	A	5.9	Met40-Cytb557	B
12	6.0	Met331	A	5.8	Ca2-CP43	B
13	5.9	Met35-PsbH	A	5.6	Cys288-CP43	A
14	5.7	Met198-D2	B	5.6	Met359-CP47	B
15	5.6	Met293	A	5.5	Met138-CP47	A
16	5.5	Cys71-D2	A	5.3	Met25-CP47	A
17	5.5	Met183	B	5.3	Met329-D2	A
18	5.5	Met330-CP47	B	5.2	CL2	B
19	5.5	Met37-CP47	B	5.2	Met342-CP43	B
20	5.5	Cys37-PsbV	A	5.1	Ca1-OEC	B
21	5.5	Met359-CP47	B	5.0	Ca-PsbO	A
22	5.3	Met194	A	5.0	Ca2-CP43	A
23	5.2	Cys44-PsbO	B	5.0	Met1-PsbT	B
24	5.2	Met198-D2	A	5.0	Met256-CP47	A
25	5.2	SQDG	A	5.0	Met271-D2	A
26	5.1	Met1-PsbT	A	4.9	Met330-CP47	A
27	5.1	Cys19-PsbO	A	4.9	Met246-D2	B
28	5.0	Met469-CP43	B	4.9	Met293	A
29	5.0	Cys47	A	4.8	Met60-CP47	B
30	5.0	Cys150-CP47	B	4.8	Met104-PsbV	A
1σ (e^- / Å^3)		0.00933			0.00988	