

Received 21 September 2020

Accepted 24 February 2021

Edited by F. Maia, Uppsala University, Sweden

‡ These authors contributed equally to this work.

§ Current address: MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK.

**Keywords:** time-resolved serial crystallography; X-ray free-electron lasers; membrane proteins; photosystem II; serial crystallography; molecular movies; protein structures.**PDB references:** dark1, 7cou; dark2, 7cjj; 10 ms, light, S<sub>2</sub> state, 7cjj**Supporting information:** this article has supporting information at [www.iucrj.org](http://www.iucrj.org)

# Capturing structural changes of the S<sub>1</sub> to S<sub>2</sub> transition of photosystem II using time-resolved serial femtosecond crystallography

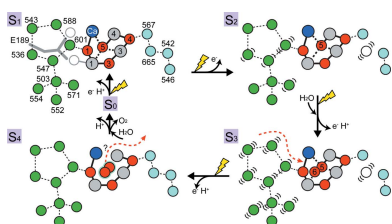
Hongjie Li,<sup>a,b,‡</sup> Yoshiki Nakajima,<sup>a,‡</sup> Takashi Nomura,<sup>b,c</sup> Michihiro Sugahara,<sup>c</sup> Shinichiro Yonekura,<sup>a</sup> Siu Kit Chan,<sup>a</sup> Takanori Nakane,<sup>d,§</sup> Takahiro Yamane,<sup>a</sup> Yasufumi Umena,<sup>a</sup> Mamoru Suzuki,<sup>e</sup> Tetsuya Masuda,<sup>f</sup> Taiki Motomura,<sup>a,b</sup> Hisashi Naitow,<sup>c</sup> Yoshinori Matsuura,<sup>c</sup> Tetsunari Kimura,<sup>g</sup> Kensuke Tono,<sup>c,h</sup> Shigeki Owada,<sup>c,h</sup> Yasumasa Joti,<sup>c,h</sup> Rie Tanaka,<sup>c,i</sup> Eriko Nango,<sup>c,j</sup> Fusamichi Akita,<sup>a,k</sup> Minoru Kubo,<sup>b,c</sup> So Iwata,<sup>c,i</sup> Jian-Ren Shen<sup>a,\*</sup> and Michihiro Suga<sup>a,k,\*</sup>

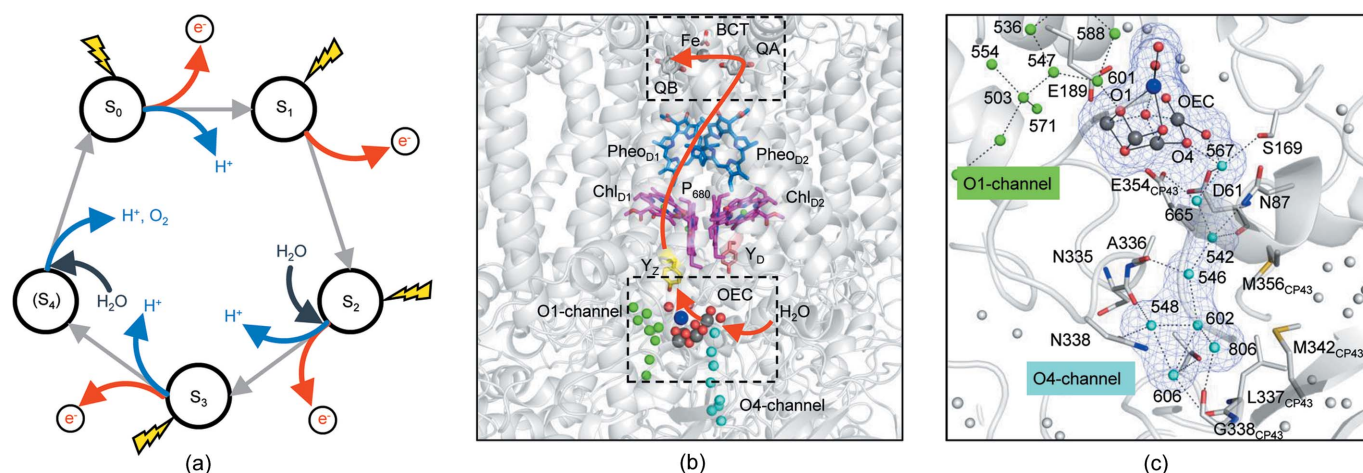
<sup>a</sup>Research Institute for Interdisciplinary Science and Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tushima-naka, Kitaku, Okayama, Okayama 700-8530, Japan, <sup>b</sup>Graduate School of Life Science, University of Hyogo, 3-2-1 Kouto, Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan, <sup>c</sup>RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan, <sup>d</sup>Department of Biological Science, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, <sup>e</sup>Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan, <sup>f</sup>Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan, <sup>g</sup>Department of Chemistry, Graduate School of Science, Kobe University, -1 Rokkodai, Nada-ku, Kobe 657-8501, Japan, <sup>h</sup>Japan Synchrotron Radiation Research Institute, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5198, Japan, <sup>i</sup>Department of Cell Biology, Graduate School of Medicine, Kyoto University, Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan, <sup>j</sup>Institute of Multi-disciplinary Research for Advanced Materials, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan, and <sup>k</sup>Japan Science and Technology Agency, PRESTO, Saitama 332-0012, Japan. \*Correspondence e-mail: [shen@cc.okayama-u.ac.jp](mailto:shen@cc.okayama-u.ac.jp), [michisuga@okayama-u.ac.jp](mailto:michisuga@okayama-u.ac.jp)

Photosystem II (PSII) catalyzes light-induced water oxidation through an S<sub>n</sub>-state cycle, leading to the generation of di-oxygen, protons and electrons. Pump-probe time-resolved serial femtosecond crystallography (TR-SFX) has been used to capture structural dynamics of light-sensitive proteins. In this approach, it is crucial to avoid light contamination in the samples when analyzing a particular reaction intermediate. Here, a method for determining a condition that avoids light contamination of the PSII microcrystals while minimizing sample consumption in TR-SFX is described. By swapping the pump and probe pulses with a very short delay between them, the structural changes that occur during the S<sub>1</sub>-to-S<sub>2</sub> transition were examined and a boundary of the excitation region was accurately determined. With the sample flow rate and concomitant illumination conditions determined, the S<sub>2</sub>-state structure of PSII could be analyzed at room temperature, revealing the structural changes that occur during the S<sub>1</sub>-to-S<sub>2</sub> transition at ambient temperature. Though the structure of the manganese cluster was similar to previous studies, the behaviors of the water molecules in the two channels (O1 and O4 channels) were found to be different. By comparing with the previous studies performed at low temperature or with a different delay time, the possible channels for water inlet and structural changes important for the water-splitting reaction were revealed.

## 1. Introduction

Oxygenic photosynthesis converts light energy into chemical energy, thereby sustaining all aerobic life on Earth. The energy-conversion reaction of photosynthesis is carried out by two photosystems (PSs), PSI and PSII, and both are large membrane-embedded protein-pigment complexes existing on the thylakoid membranes of plants and various algae. PSII uses light energy to extract electrons and protons from water molecules, leading to the oxidation of water and the release of dioxygen as a byproduct. This water-oxidation reaction is catalyzed by the oxygen-evolving complex (OEC) of PSII and





**Figure 1**

(a)  $S_i$ -state cycle of the water-oxidation reaction of the OEC. (b) Electron-transfer chain of PSII. The flow of electrons is indicated by a red line, and the regions around the OEC,  $Q_B$  and the non-heme iron are boxed with black dashed lines. (c) Enlarged views of the boxed region around the OEC shown in (b). Water molecules in the O1 and O4 channels are shown as green and cyan spheres, respectively. The light-blue mesh shows the radius of atoms of the OEC and the water molecules in the O4 channel. The Cl1 and Cl2 channels have been omitted for clarity.

proceeds through a light-driven, five-step  $S_i$ -state cycle ( $i = 0-4$ ) (Kok *et al.*, 1970; Shen, 2015; Cox *et al.*, 2020) [Fig. 1(a)]. In this  $S_i$ -state cycle,  $S_0$  is the ground state, and the OEC progresses to higher  $S_i$  states upon its oxidation by  $Y_Z^+$ , a tyrosine cation residue generated by  $P_{680}^+$  the photoexcited PSII reaction center  $P_{680}$  [Fig. 1(b)]. The release of dioxygen occurs during the  $S_3$ -to- $(S_4)$ -to- $S_0$  transition. The OEC is dark-stable in the  $S_1$  state, with the chemical composition  $Mn_4CaO_5$  (Umena *et al.*, 2011; Suga *et al.*, 2015). The OEC changes to an  $Mn_4CaO_6$  cluster in the highest metastable  $S_3$  state by incorporation of a new oxygen atom (O6) near the unique  $\mu$ -oxo bridge O5, which was first identified at 2.35 Å resolution (Suga *et al.*, 2017), and further analyzed at resolutions of 2.15 Å (Suga *et al.*, 2019) and 2.09 Å [the new oxygen atom was referred to as Ox (Ibrahim *et al.*, 2020)]. The interatomic distance between O6 and O5 has been reported to be 1.9 Å (Suga *et al.*, 2019) or 2.2 Å (Ibrahim *et al.*, 2020), which are suitable for an oxyl/oxo type coupling to form dioxygen between them. The detailed mechanism of the water oxidation, however, has not been well understood, in particular concerning the origin of O6 and the proton exit pathways.

The  $Mn_4CaO_5$  cluster is embedded inside the protein matrix of PSII and covered by a large area of hydrophilic protein regions in the luminal side of the thylakoid membrane. Channels for the inlet of substrate waters and egress of the product protons are important for the water-splitting reaction to proceed appropriately. The high-resolution structure of PSII showed multiple hydrogen-bonded networks connecting the site of the  $Mn_4CaO_5$  cluster to the luminal surface of PSII (Umena *et al.*, 2011; Shen, 2015; Suga *et al.*, 2015). These channels may therefore function to allow water into the catalytic site or protons to be transported to the lumen. Four main such channels have been identified: the O1 channel, O4 channel, Cl1 channel and Cl2 channel (Fig. 1). The O1 and O4 channels are so-called because they start from the oxo-bridges O1 and O4 of the  $Mn_4CaO_5$  cluster, respectively, whereas the Cl1 and Cl2 channels are mediated by the Cl1 and Cl2 ions in

the vicinity of the  $Mn_4CaO_5$  cluster, respectively. It is not clear which of these channels functions in the water inlet or proton egress, and in the latter case, in which  $S$ -state transition.

Pump-probe, time-resolved serial femtosecond X-ray crystallography (TR-SFX) using X-ray free-electron lasers (XFEL) is a powerful method to visualize structural dynamics of light-sensitive proteins (Tenboer *et al.*, 2014; Barends *et al.*, 2015; Nango *et al.*, 2016; Pande *et al.*, 2016; Nogly *et al.*, 2018; Suga *et al.*, 2020), including PSII in different  $S_i$  states at ambient temperature (Kern *et al.*, 2013, 2014, 2018; Kupitz *et al.*, 2014; Young *et al.*, 2016; Suga *et al.*, 2017; Ibrahim *et al.*, 2020) or cryogenic temperature (Suga *et al.*, 2019). XFELs provide femtosecond pulses of X-rays with an approximately billion-fold increase in peak brilliance when compared with conventional synchrotron X-rays, thus enabling collection of diffraction data before the onset of radiation damage (Neutze *et al.*, 2000). For capturing the intermediate  $S_i$  states of PSII with this method, a flow of PSII microcrystals either in solution or embedded in a matrix is illuminated by a desired number of pump flashes to generate the higher  $S_i$  states (one, two or three flashes generate  $S_2$ ,  $S_3$  or  $S_0$  states, respectively), followed by detection with an XFEL pulse with a temporal delay time after the flash illumination. One of the most critical factors for the success of this experiment is the selection of an optimal light excitation condition (intervals, the boundary of excitation region and power of the excitation laser *etc.*) under particular sample delivery conditions (crystal size, flow diameter, flow rate, overall sample consumption *etc.*). While the light illumination area is primarily defined by the laser spot (Kovacs *et al.*, 2019; Grünbein *et al.*, 2020), some reports claimed that the light may scatter to a larger area than its illumination spot (Nogly *et al.*, 2018). A generally applicable way to determine an effective light excitation condition is adding a generous safety margin in the sample area to be excited and examining the structural changes by TR-SFX under a suitable size and power of the pump laser illumination. However, in the case of enzymes such as PSII that

requires multi-flash excitations for the higher S-states, the application of this method is not straightforward under the continuous sample flow condition. The sample area that can be used for the excitation and X-ray diffraction must be well aligned spatially, and the pump/probe illumination area and the sample must be aligned temporally. However, a larger separation between two consecutive flashes may cause the illuminated sample to escape from the area under irradiation by the XFEL pulses. In such cases, higher S<sub>i</sub> states may not be captured by this TR-SFX method, and other methods such as fixed-target crystallography (Suga *et al.*, 2019) or a tape drive (Ibrahim *et al.*, 2020) may be potentially applicable.

In the present study, we developed a method to determine an optimal light illumination condition for successful TR-SFX to analyze the structures of the intermediate S<sub>i</sub> states of PSII. By altering the flash interval distances with maximum delay time and examining the structural changes that occur during the S<sub>1</sub>-to-S<sub>2</sub> transition, a boundary of the excitation region was determined. Based on the light illumination conditions determined, we analyzed the PSII structure in the S<sub>2</sub> state at 2.4 Å resolution. Structural changes were found in the OEC, the O1 and O4 channels, and the Q<sub>B</sub>-binding site, providing important insights into the substrate water delivery during the water-oxidation reaction.

## 2. Methods

### 2.1. Preparation and crystallization of photosystem II

Highly active dimeric PSII core complexes were purified from the thermophilic cyanobacterium *Thermosynechococcus vulcanus* cells and crystallized as described previously (Shen & Kamiya, 2000; Umena *et al.*, 2011; Suga *et al.*, 2017). The microcrystals grew to a maximum size of 100 µm. They showed good quality in diffraction as well as high efficiency in the progress of the S<sub>i</sub>-state cycle (Kato *et al.*, 2018). Before the TR-SFX experiments, the microcrystals were pre-flashed by an Nd:YAG laser at 532 nm with a diameter of 7 mm at an energy of around 52 mJ cm<sup>-1</sup> (Suga *et al.*, 2017) to oxidize the tyrosine D residue (D2-Y160) and decrease the contamination of the S<sub>0</sub> state. The microcrystals were dehydrated with a buffer containing 20% glycerol, 10% PEG 1450, 10% PEG 5000 MME, 2% dimethyl sulfoxide (DMSO) and 10 mM potassium ferricyanide by stepwise replacement of the solution, which took around 2 h. Following the dehydration, the crystals were mixed with a silicon grease (Sugahara *et al.*, 2015) and used for the TR-SFX experiments. All the procedures, including purification, crystallization, pre-flash illumination, dehydration and mixing with the grease matrix were performed in dark- or dim-green light.

### 2.2. Diffraction experiments at SACLA-XFEL

The PSII crystals mixed with the grease matrix were loaded into an injector (Shimazu *et al.*, 2019) with a nozzle diameter of 140 µm and set in a pump-probe system based on Diverse Application Platform for Hard X-ray Diffraction (DAPHNIS) in the SPring-8 Ångstrom Compact Free Electron Laser

(SACLA) (Tono *et al.*, 2015; Kubo *et al.*, 2017). The flow rates were as follows: 2.5 µl min<sup>-1</sup> for the dark datasets; 4.9, 7.3, 8.5 and 9.8 µl min<sup>-1</sup> for each ‘light’ dataset at a delay time of -50 ns (that is, 50 ns before the laser flash); 9.8 µl min<sup>-1</sup> for the light dataset at a delay time of 10 ms. The actual speed of the stream (the distance traveled by the sample per unit time) under the volumetric flow rate specified was examined at the preliminary stages of the experiment. We monitored the stream of the PSII microcrystals using a high-speed camera at a flow rate of 7.8 µl min<sup>-1</sup> with a nozzle diameter of 125 µm. Under this condition, the speed of the stream was expected to be 10.6 mm s<sup>-1</sup>, whereas the actual speed and diameter of the stream were 9.7 ± 0.9 mm s<sup>-1</sup> and 131 ± 7 µm, respectively. When the stream was stable, the actual speed was constant. This was also confirmed with microcrystals in the lipid cubic phase (Nango *et al.*, 2019). Although we did not monitor the speed of the stream throughout the experiment, we collected diffraction images only when the stream was stable, and we stopped data collection when the stream showed a stop/go behavior or balled up for a limited amount of time, or wiggled *etc.*

As the pump laser, a 532 nm pulse from an Nd:YAG laser source (Minilite-I, Continuum) with a repetition of 10 Hz, was split into two beams, and each beam was focused on the microcrystals from two different directions with an angle of 160° for efficient excitation (Suga *et al.*, 2017). The pump focal diameter was set to 240 µm at the targeted sample position, and its energy was 42 mJ cm<sup>-2</sup> from each direction. Diffraction images were collected using femtosecond X-ray pulses from SACLA at BL3 with the following pulse parameters: 2–10 fs pulse duration, 7 keV X-ray energy, 0.5% (FWHM) energy bandwidth, pulse flux: ~7 × 10<sup>10</sup> photons per pulse, 3.0 (H) × 3.0 µm (W) beam size, 30 Hz repetition rate. The XFEL pulses were provided to the pump focal center either before 50 ns of the pump laser pulse for the ‘light’ datasets at different flow rates, or 106 µm downstream from the pump focal center after 10 ms of the pump laser pulse for the light dataset, which fully transforms the S state to the S<sub>2</sub> state. Diffraction images were recorded by a multiport CCD detector. Because the excitation laser pulses were provided at 10 Hz and the XFEL pulses had a repetition rate of 30 Hz, all ‘pump-on’ images were recorded at 10 Hz for the ‘light’ datasets, whereas the diffraction data for the dark dataset were recorded at 30 Hz in a separate run.

### 2.3. Data processing

The diffraction data collected at SACLA were monitored by the program *Cheetah* (Barty *et al.*, 2014; Nakane *et al.*, 2016), and the diffraction images passing through the ‘filter’ with pre-defined thresholds of diffraction spot numbers (recorded as ‘hits’) were used for the subsequent processing. Indexing, integration, scaling and merging of the images were carried out by programs in the *cctbx.xfel* suite (Sauter *et al.*, 2013; Sauter, 2015; Waterman *et al.*, 2016). The diffraction images were indexed and integrated by the program *dials.stills\_process* using the geometry of the detector and a

**Table 1**

Statistics for data collection and structure refinement.

Values in parentheses are those of the highest resolution shell. Structure factors for  $-50$  ns, light datasets at flow rates of  $4.9 \mu\text{l min}^{-1}$  (Asf),  $7.3 \mu\text{l min}^{-1}$  (Bsf),  $8.5 \mu\text{l min}^{-1}$  (Csf), and  $9.8 \mu\text{l min}^{-1}$  (Dsf) have been deposited in the Protein Data Bank (PDB) with the accession number 7cou.

Data name	Dark1	Dark2	$-50$ ns, light, $4.9 \mu\text{l min}^{-1}$	$-50$ ns, light, $7.3 \mu\text{l min}^{-1}$	$-50$ ns, light, $8.5 \mu\text{l min}^{-1}$	$-50$ ns, light, $9.8 \mu\text{l min}^{-1}$	10 ms, light, $9.8 \mu\text{l min}^{-1}$
Data collection							
Flow rate ( $\mu\text{l min}^{-1}$ )	2.5	2.5	4.9	7.3	8.5	9.8	9.8
No. of indexed images	96459	41071	18216	13858	20449	22419	17247
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 ( $\text{\AA}$ )	$a = 126.1$ $b = 231.7$ $c = 288.3$	$a = 126.0$ $b = 231.7$ $c = 288.3$	$a = 125.8$ $b = 231.5$ $c = 288.2$	$a = 126.0$ $b = 231.6$ $c = 288.2$	$a = 125.7$ $b = 231.6$ $c = 288.4$	$a = 125.6$ $b = 231.7$ $c = 288.5$	$a = 125.7$ $b = 231.5$ $c = 288.3$
Resolution ( $\text{\AA}$ )	40–2.25	40–2.35	40–2.40	40–2.40	40–2.40	40–2.40	40–2.40
Highest shell ( $\text{\AA}$ )	2.33–2.25	2.43–2.35	2.49–2.40	2.49–2.40	2.49–2.40	2.49–2.40	2.49–2.40
No. of unique reflections	399661	351162	329839	329852	329843	329840	329844
Completeness (%)	100	100	100	100	100	100	100
Multiplicity	847 (547)	494 (339)	159 (109)	111 (76)	249 (172)	271 (187)	235 (162)
$R_{\text{split}}$ (%)	4.6 (52.4)	5.7 (56.5)	8.1 (60.7)	10.8 (76.6)	6.7 (53.3)	7.3 (56.4)	8.4 (63.8)
$CC_{1/2}$	0.999 (0.62)	0.999 (0.60)	0.996 (0.610)	0.994 (0.452)	0.998 (0.706)	0.997 (0.643)	0.997 (0.540)
Mean $I/\sigma(I)$	65.3 (2.2)	56.8 (2.0)	36.9 (1.8)	27.1 (1.4)	43.1 (2.1)	40.2 (2.0)	37.0 (1.8)
Refinement							
$R_{\text{work}}/R_{\text{free}}$ (%)	0.169/0.211	0.160/0.208	–	–	–	–	0.168/0.219
Wilson $B$ ( $\text{\AA}^2$ )	44.0	48.4	–	–	–	–	46.1
Average $B$ factor ( $\text{\AA}^2$ )	61.3	67.1	–	–	–	–	63.1
Protein	62.2	67.0	–	–	–	–	63.0
OEC	43.2	51.2	–	–	–	–	48.1
Water	64.0	69.1	–	–	–	–	66.0
RMSD bond length ( $\text{\AA}$ )	0.008	0.008	–	–	–	–	0.008
RMSD bond angle ( $^\circ$ )	1.238	1.245	–	–	–	–	1.263
Ramachandran (%)							
Favoured	97.71	97.64	–	–	–	–	97.66
Allowed	2.11	2.25	–	–	–	–	2.21
Outliers	0.17	0.12	–	–	–	–	0.13
PDB entry	7cou	7cji	7cou	7cou	7cou	7cou	7cjj

camera distance refined by the program *cspad.cbf\_metrology* (Brewster *et al.*, 2018). The unit-cell parameters ( $a = 126.3$ ,  $b = 232.1$ ,  $c = 289.0 \text{ \AA}$ , and  $\alpha = \beta = \gamma = 90^\circ$ ), originally determined by the program *CrystFEL* (White *et al.*, 2012) were provided during the indexing process. The number of indexed images for each dataset were as follows: 96459 for dark1; 41071 for dark2; 18216 for  $-50$  ns ‘light’ at  $4.9 \mu\text{l min}^{-1}$ ; 13858 for  $-50$  ns ‘light’ at  $7.3 \mu\text{l min}^{-1}$ ; 20449 for  $-50$  ns ‘light’ at  $8.5 \mu\text{l min}^{-1}$ ; 22419 for  $-50$  ns ‘light’ at  $9.8 \mu\text{l min}^{-1}$ ; 17247 for 10 ms light at  $9.8 \mu\text{l min}^{-1}$  (Table 1). Integrated images were merged by the program *cxi.merge* with the post-refinement rs2 algorithm, and a filter based on the value of  $I/\sigma(I)$  was not applied in order to include weak signals at high resolutions. Dataset dark1 or dark2 could be used as the reference model in the processing of ‘light’ datasets from the same purification batch of samples. All datasets were processed to 2.25–2.40  $\text{\AA}$  resolutions based on the criteria of  $CC_{1/2} < 50\%$ . The  $R_{\text{iso}}$  values between the datasets were 0.073–0.100 in the same sample batch, and 0.095–0.112 between the different sample batches. Because different sample batches gave a slightly higher  $R_{\text{iso}}$  between the different datasets, and because a lower  $R_{\text{iso}}$  was crucial to detect meaningful structural changes based on the isomorphous difference Fourier map, we collected the two different dark datasets (dark1 and dark2) independently for the different sample batches used.

#### 2.4. Model building and map calculation

The initial phase of the dark dataset was obtained by molecular replacement with the program *Phaser-MR* in the *CCP4* package (Collaborative Computational Project, 1994), using the previous PSII structure at room temperature (PDB entry 5ws5; Suga *et al.*, 2017) as the search model. Then, an initial rigid-body refinement, followed by refinement of coordinates,  $B$  factors and TLS, was performed using the program *Phenix* (Adams *et al.*, 2010; Afonine *et al.*, 2012) combined with manual modifications of the program *Coot* (Emsley & Cowtan, 2004). The restraints used here were the same as those used previously (Suga *et al.*, 2017, 2019). The final  $R_{\text{work}}$  and  $R_{\text{free}}$  were 0.169 and 0.211 for the dark1 dataset, and 0.160 and 0.208 for the dark2 dataset (Table 1). The phases from the refined model of dark1 or dark2 were used to calculate the isomorphous difference Fourier maps between dark and each  $-50$  ns, ‘light’ dataset, or 10 ms, light dataset. Strong peaks in the isomorphous difference Fourier maps were found around the regions of the OEC,  $Q_B$  and non-heme iron, and these regions were built as a mixture of dominant  $S_2$  and minor  $S_1$  states based on the transition efficiency of the PSII microcrystals (Kato *et al.*, 2018). We first refined the coordinates of the dominant  $S_2$  state only with fixed  $B$  factors. Then we refined the  $B$  factors of atoms in the regions mentioned above which show structural changes during the



$S_1$ -to- $S_2$  transition by assuming the populations of the  $S_2$  and  $S_1$  states at (0.9, 0.1), (0.8, 0.2), (0.7, 0.3) and (0.6, 0.4), and found that the residual electron densities and  $B$  factors after the refinement were reasonable for the populations of  $S_2$  and  $S_1$  states of 0.7 and 0.3, respectively. The determined occupancy was consistent with the transition efficiency of the PSII microcrystals (Kato *et al.*, 2018). Thus, this occupancy was used to refine the final structure in the  $S_2$  state. This includes 29 waters, 42 residues, 2 OECs, 2  $Q_{BS}$ , 2 BCTs and 2 non-heme irons. The statistics for structural refinement are provided in Table 1.

### 3. Results

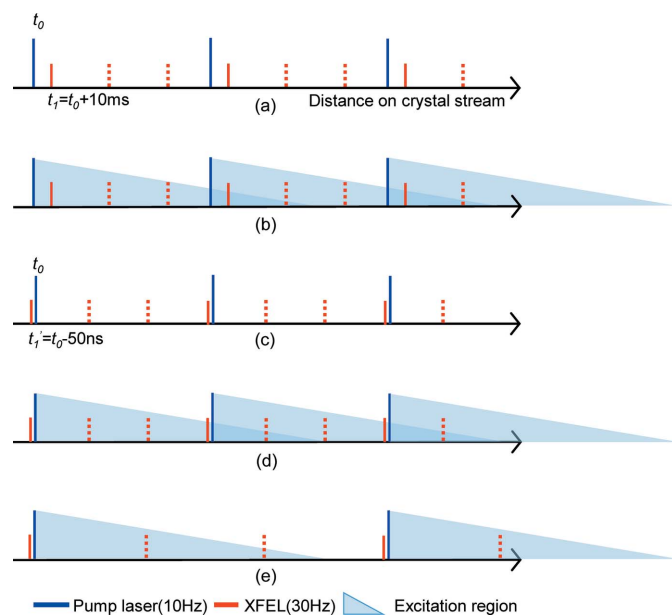
#### 3.1. Determination of a boundary of the excitation region using TR-SFX

We performed TR-SFX as described previously (Suga *et al.*, 2017; Sugahara *et al.*, 2015). In this approach, PSII microcrystals were mixed with a grease matrix and ejected from a micro-extrusion injector. The flow of the PSII microcrystals was excited by a single flash to advance the S state to  $S_2$ . Fig. 2(a) shows a scheme representing the interaction between the pump excitation region and the XFEL pulse in the TR-SFX experiment with a delay time ( $\Delta t$ ) of 10 ms at repetition rates of 10 and 30 Hz for the pump and XFEL, respectively. It should be noted that, although the pump beam was focused on the sample stream with a top-hat shape ( $\phi = 250 \mu\text{m}$ ) (Kubo *et*

*al.*, 2017), the effective excitation region extended upstream and downstream possibly due to the pump-light scattering on the stream as well as tailing of the laser illumination area. Thus, the excitation region is schematically depicted by a triangle in the figure, indicating a gradual decrease in the pump photon density along the sample stream. When a flow rate is not fast enough for the sample exchange, the pump excitation region interacts with the next pump-XFEL pulse [Fig. 2(b)]. To avoid such erroneous light-contamination, we designed a test experiment with  $\Delta t = -50 \text{ ns}$ , where the negative time points signify that the X-rays arrive 50 ns before the laser, so that the delay to the next laser pulse (at 10 Hz) is 99.99995 ms for the ‘pump-on’ data [Fig. 2(c)]. As the XFEL was operated at 30 Hz, there will also be ‘pump-off’ data between the two ‘pump-on’ data [red dashes in Fig. 2(c)], which were not processed in the present study. On this  $-\Delta t$  condition, if the flow rate is slow, the microcrystals would be illuminated partially by the preceding flash, resulting in ‘one-flash’ illumination [Fig. 2(d)]. However, at a sufficiently fast flow rate, the microcrystals at the position of the XFEL pulse will escape from the preceding flash illumination [Fig. 2(e)], resulting in a ‘dark dataset.’ Accordingly, we can check light contamination, including a possible effect of pump light scattering on the sample stream, under the given experimental conditions (pump-illumination size and intensity, sample flow rate *etc.*).

Four diffraction datasets were collected at different flow rates of 4.9, 7.3, 8.5 and 9.8  $\mu\text{l min}^{-1}$  (corresponding to 2.0 times, 3.0 times, 3.5 times and 4.0 times that for the dark datasets, respectively). In addition to these four datasets, we collected two independent dark datasets by different preparations (dark1 for the light-illuminated, flow rate 4.9 and 7.3  $\mu\text{l min}^{-1}$  experiments, and dark2 for the light-illuminated, flow rate 8.5 and 9.8  $\mu\text{l min}^{-1}$  experiments) at a flow rate of 2.5  $\mu\text{l min}^{-1}$ , and a light dataset with  $\Delta t = 10 \text{ ms}$  at a flow rate of 9.8  $\mu\text{l min}^{-1}$ . All datasets were processed at 2.25 to 2.40 Å resolutions (Table 1).

We evaluated the boundary of the excitation region as follows. Isomorphous difference Fourier maps between the ‘light’-illuminated and dark datasets were calculated with the phases obtained by the refinement of the dark datasets, which showed a negative peak covering W665, the second water molecule from O4 in the O4 channel [Fig. 1(c)]. Since a similar negative peak has been observed in the previous studies, indicating that W665 becomes highly disordered during the  $S_1$ -to- $S_2$  transition (Suga *et al.*, 2019; Kern *et al.*, 2018), we take the height of the Fourier difference peak of W665 as an indicator for the light-induced structural changes. The  $F_{\text{obs}}$  ( $-50 \text{ ns}$ , ‘light’ at 4.9  $\mu\text{l min}^{-1}$ ) minus  $F_{\text{obs}}$  (dark) difference Fourier map showed a peak height of  $-6.9\sigma$  at the position of W665 [Fig. 3(a), Table S1 of the supporting information]. This peak height is lower than that observed with a delay time of 10 ms after the excitation flash [Fig. 3(e)], but is higher than the maximum noise level or systematic errors [Fig. 3(f), Table S1], suggesting that, at this flow rate, the microcrystals at the target position of the XFEL shot have been excited by the preceding flash illumination. Thus, the flow rate of

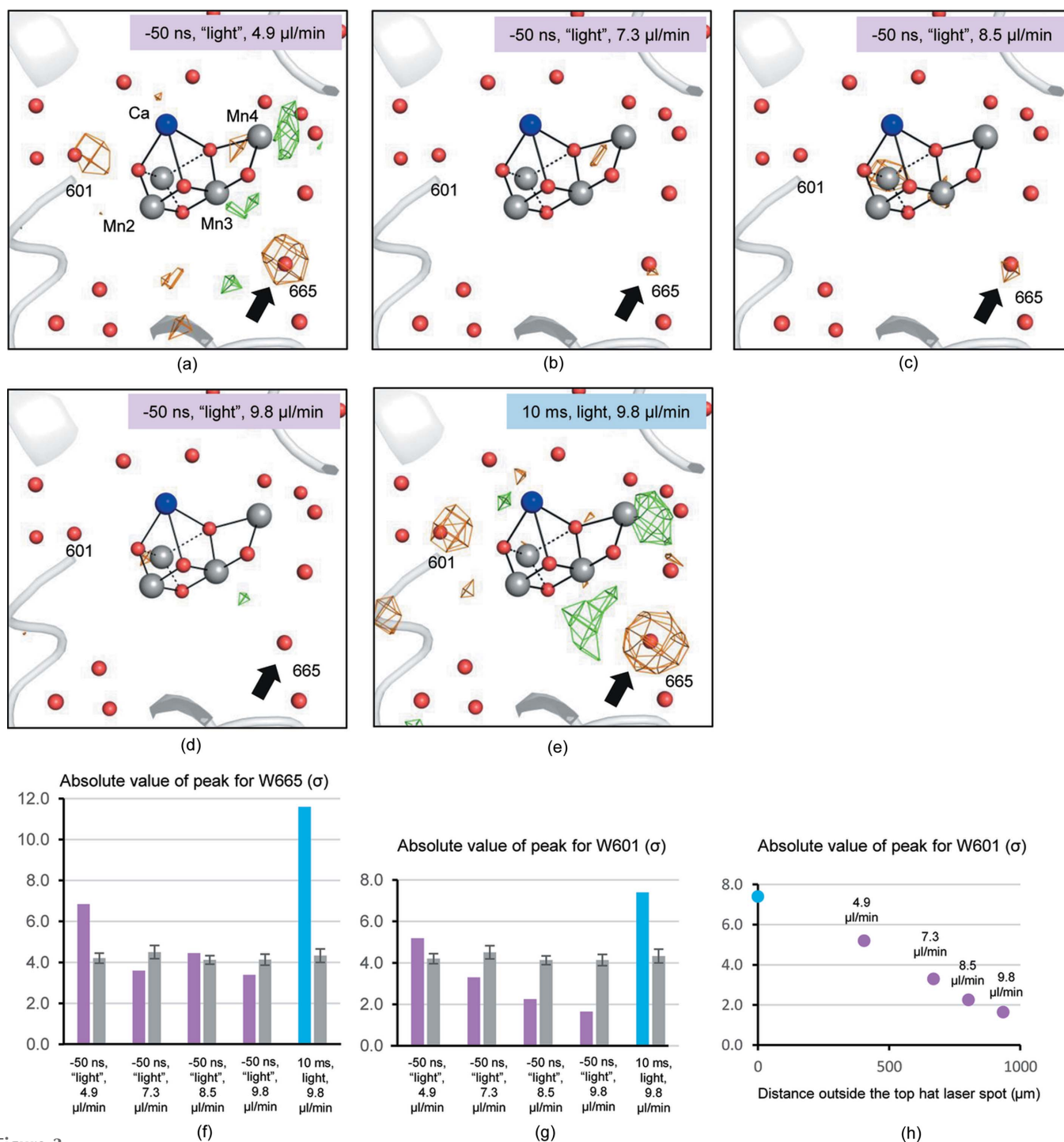


**Figure 2**

Relative timing of the pump lasers and XFEL pulses with  $\Delta t = 10 \text{ ms}$  (a) and  $\Delta t = -50 \text{ ns}$  (c). Schematic representations of the boundaries where the pump lasers reaching a slower flow rate with  $\Delta t = 10 \text{ ms}$  (b) or  $\Delta t = -50 \text{ ns}$  (d), and at a faster flow rate with  $\Delta t = -50 \text{ ns}$  (e). Note that the region exposed to XFEL pulses interacts with multiple pump lasers (b) or an unintended laser (d) when a slower flow rate was employed. XFEL pulses that were not recorded for the ‘light’ datasets are shown as dashed lines. The excitation region is shown as a triangle for clarity, but it is not a linear profile in the experiment.

$4.9 \mu\text{ l min}^{-1}$  is too slow to avoid light-contamination by the preceding flash in the excitation region. The average height of the light-minus-dark Fourier difference peak of W665 in two non-crystallographic symmetry-related PSII monomers was

reduced when the flow rate was increased [Figs. 3(b)–3(d), Table S1], and reached a level not visible in the difference map contoured at  $\pm 4.0\sigma$  at a flow rate of  $9.8 \mu\text{ l min}^{-1}$  [Fig. 3(d)], which is also well below the maximum noise level or



**Figure 3** (a)–(e)  $S_1$ -state structure superimposed with the  $F_{\text{obs}}(\text{light}) - F_{\text{obs}}(\text{dark})$  isomorphous difference Fourier map contoured at  $+4.0\sigma$  (green) and  $-4.0\sigma$  (orange) calculated with the ‘light’ datasets obtained under the conditions (a) –50 ns, ‘light’,  $4.9 \mu\text{ l min}^{-1}$  flow rate; (b) –50 ns, ‘light’,  $7.3 \mu\text{ l min}^{-1}$ ; (c) –50 ns, ‘light’,  $8.5 \mu\text{ l min}^{-1}$ , (d) –50 ns, ‘light’,  $9.8 \mu\text{ l min}^{-1}$ ; and (e) 10 ms, light,  $9.8 \mu\text{ l min}^{-1}$ . The changes in W665 were indicated by black arrows. Average peak heights of the difference map at the position of (f) W665 and (g) W601 calculated from two PSII monomers and maximum noises or systematic errors. Maximum noise or systematic errors with error bars were calculated from the five strongest noise peaks observed outside the PSII protein complex and are shown in gray. (h) Peak heights at the position of W601 were plotted against the distance the sample traveled outside the top-hat laser spot where the sample is no longer directly illuminated. The 10 ms, light,  $9.8 \mu\text{ l min}^{-1}$  dataset corresponds to distance zero (the focus of the laser beam), and for the other datasets with a delay time of –50 ns, the calculated total distance traveled –125  $\mu\text{ m}$  was used.

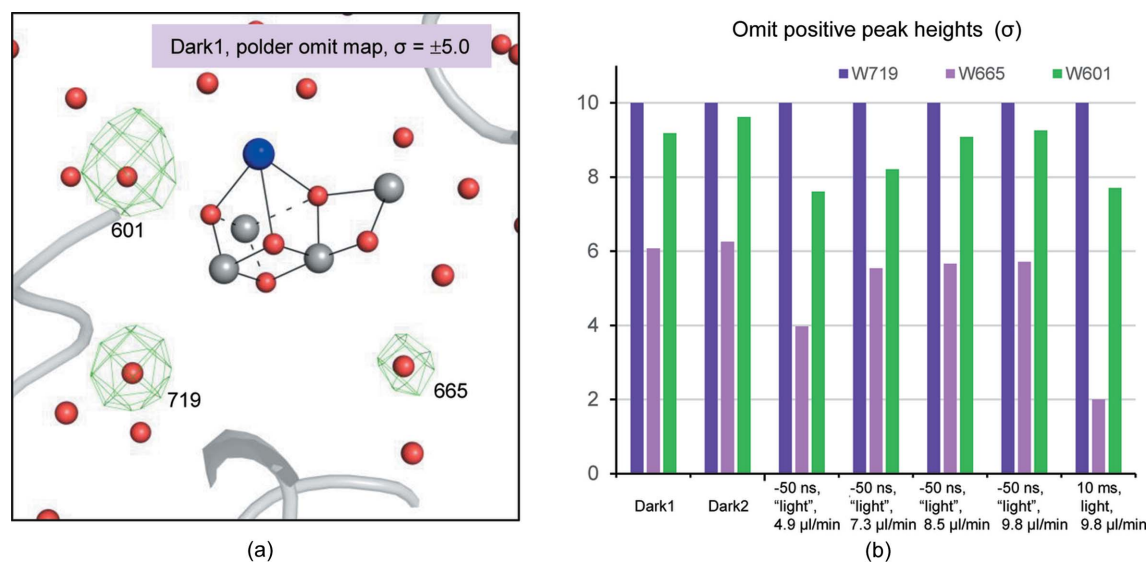
systematic errors [Fig. 3(f)]. This indicates that at the flow rate  $9.8 \mu\text{L min}^{-1}$ , no apparent light-minus-dark Fourier difference peak was observed. Unexpectedly, the peak at the  $8.5 \mu\text{L min}^{-1}$  flow rate ( $-4.5\sigma$ ) was slightly higher than that of  $7.3 \mu\text{L min}^{-1}$  ( $-3.6\sigma$ ) (Table S1). So we also compared the peak heights of W601 (another water molecule that disappeared in the  $S_1$ -to- $S_2$  transition, discussed in more detail below) and confirmed that the signal decreased consistently [Fig. 3(g), Table S1]. These results suggest that the variations in the peak heights of W665, as well as many other changes (data not shown) in the structure derived from data collected at flow rates of 7.3 and  $8.5 \mu\text{L min}^{-1}$  may originate from non-isomorphism between the different crystal batches.

The excitation region was further estimated based on difference densities of water molecules, as the nozzle diameter used to collect the data was  $140 \mu\text{m}$ , which was different from that of  $125 \mu\text{m}$  used for the off-line monitoring of the flow rate (see Methods), suggesting an error of about 10% in the real flow rate when compared with that calculated. When the peak height of W601 was plotted with respect to the distance the sample traveled outside the top-hat laser spot where the sample is no longer directly illuminated, a high excitation yield was observed even at  $405 \mu\text{m}$ , but was diminished at  $935 \mu\text{m}$  [Fig. 3(h)]. Because these changes are based on the isomorphous differences which are readily affected by non-isomorphism between datasets, we examined the changes based on the  $F_{\text{obs}} - F_{\text{calc}}$  omit maps. Three water molecules, W665, W601 and W719 were omitted in both PSII monomers of the dimer and the omit electron densities were scaled to the dark1 dataset [Fig. 4(b)]. Here, W719 is a well defined molecule that does not change its position upon light illumination, and therefore its positive omit electron density is used as a baseline for comparison to determine the peak heights. When compared with the dark1 and dark2 datasets, electron densi-

ties for W665 and W601 became lower at flow rates of 4.9 and  $7.3 \mu\text{L min}^{-1}$ , but were comparable at flow rates of 8.5 and  $9.8 \mu\text{L min}^{-1}$  (Fig. 4). Therefore, we concluded that the flow rate  $9.8 \mu\text{L min}^{-1}$  gives rise to no light contamination at the position of the XFEL shot. We collected the light-illuminated dataset with  $\Delta t = 10 \text{ ms}$  after the flash illumination at this flow rate to analyze the PSII structure in the  $S_2$  state.

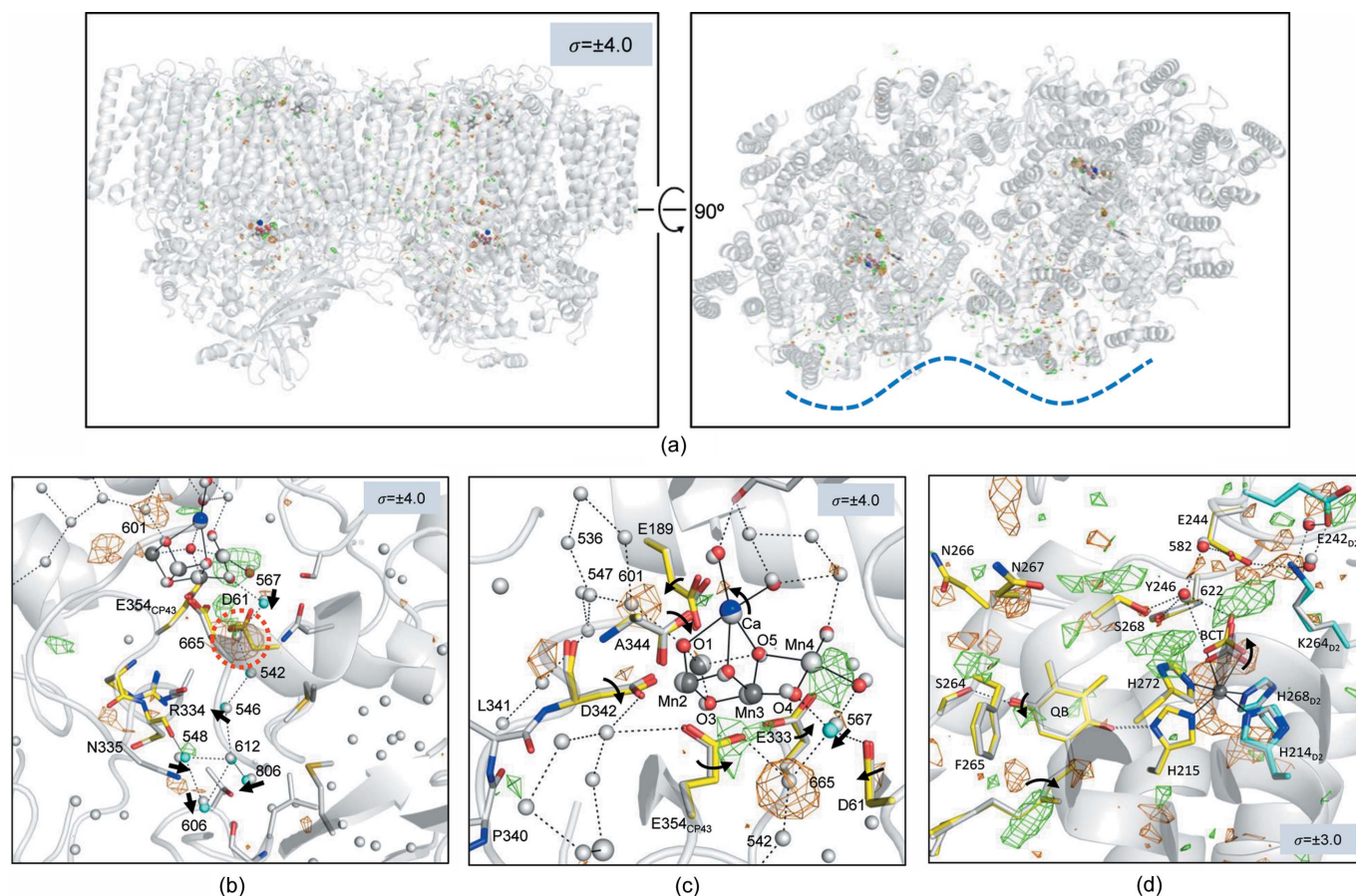
### 3.2. Structural determination of PSII in the $S_2$ state

The  $F_{\text{obs}}(10 \text{ ms, light}, 9.8 \mu\text{L min}^{-1}) - F_{\text{obs}}(\text{dark})$  isomorphous difference Fourier map showed a strong signal with a peak height above  $-11.6\sigma$  at the position of W665, suggesting that the microcrystals were successfully excited to progress to the  $S_2$  state [Figs. 5(a)–5(c)]. Many peaks can be seen in the Fourier difference map when we decrease the contour level below  $\pm 3\sigma$ ; thus we consider that the average noise level is around  $\pm 3\sigma$ , and peaks above  $\pm 3\sigma$  may represent real structural changes induced by one-flash illumination. We observe many peaks above  $\pm 4.5\sigma$  that are centered around the OEC and the electron-transfer chain, and are consistent with previously observed structural changes that occur during the  $S_1$ -to- $S_2$  transition. They are distributed around the OEC, bicarbonate (BCT), non-heme iron and  $Q_B$ , and are observed similarly in both PSII monomers. For these reasons they were interpreted as light-induced structural changes. However, some weaker peaks at around  $\pm 4\sigma$  were found at one side of the monomer–monomer interface [Fig. 5(a)]. As the isomorphism between the two datasets was relatively low, these weaker peaks may not be related to the light-induced structural changes. Instead, these may arise from differences in the batches of samples used, since it was difficult to control the sample purification conditions and sample states, such as dehydration of the crystals, to be entirely uniform.



**Figure 4**  
(a)  $S_1$ -state structure superposed with the omit map for W601, W665 and W719 calculated from the dark1 dataset contoured at  $+5.0\sigma$  (green) and  $-5.0\sigma$  (orange). (b) Positive peak heights of the omit map determined for the water molecules. W719 is a well defined water molecule that does not change its structure during the  $S$ -state cycle, thus its positive omit density was used as a baseline. The figure shows relative heights of omit maps of water molecules based on that of W719.





**Figure 5**

Structures of PSII in the  $S_1$  (gray) and  $S_2$  states (colored) superimposed with an isomorphous difference Fourier map showing (a) the overall PSII structure, and the regions of the (b) O4 channel, (c) OEC and (d)  $Q_B$  site. The difference Fourier map was contoured at  $\pm 4\sigma$  (a)–(c) and  $\pm 3\sigma$  (d) in the same color as those in Fig. 3. The surface area where the minor structural changes are distributed is indicated by a blue dashed line. Hydrogen-bonded networks of water molecules are represented by black dotted lines, and the structural changes are indicated by black arrows. W665 is encircled with a red dashed line in (b).

We refined the 10 ms light dataset as a mixture model consisting of 70%  $S_2$  state and 30%  $S_1$  state for the  $Mn_4CaO_5$  cluster, its nearby residues, the region around  $Q_B$  and the BCT binding site. These include 29 waters, 42 residues, 2 OECs, 2  $Q_B$ s, 2 BCTs and 2 non-heme irons. This gives rise to equivalent values of the temperature factors between the two equivalent atoms in the multiple model, and this population of  $S_2$  and  $S_1$  states after one-flash illumination is similar to the efficiencies of the  $S_i$  state transition estimated by Fourier transform infrared spectroscopy (Kato *et al.*, 2018; Suga *et al.*, 2017). The OEC structures in the  $S_1$  and  $S_2$  states determined in the present study are similar to those reported in previous studies (Suga *et al.*, 2019; Ibrahim *et al.*, 2020). The OEC structure in the  $S_2$  state was in the open-cubane form, in which the right side of the O5 is open, giving rise to the five-coordinate trigonal bipyramidal coordination of Mn1. All changes in the Mn–Mn and Mn–Ca distances during the  $S_1$ -to- $S_2$  transition were less than the error range of the coordinates at the current resolution (Fig. 6). However, the changes found in the isomorphous difference Fourier map, such as the shortening of Mn3–Mn4 and elongation of Mn1–Mn3, Mn3–Ca and Mn4–Ca, were consistent with the previous study with the

diffraction data collected at 100 K for the room-temperature-trapped  $S_2$  state (Suga *et al.*, 2019) (Fig. 6 and Table 2). In association with the movements of these manganese atoms, some ligand residues of the OEC (D1–E189, E333, D342, A344 and CP43–E354) also moved slightly, as previously observed (Suga *et al.*, 2019) [Fig. 5(c)].

### 3.3. Structural changes in the O4 and O1 channels

Among the four channels, the O4 channel has been suggested to function as the pathway of proton release during the  $S_0$ -to- $S_1$  transition based on theoretical calculations (Saito *et al.*, 2015). On the other hand, other groups have argued that it serves as the source of substrate water by a ‘pivot’ or ‘carousel’ mechanism in the  $S_2$ -to- $S_3$  transition (Wang *et al.*, 2017; Retegan *et al.*, 2016; Kawashima *et al.*, 2018). As described above, upon transition to the  $S_2$  state, W665 in the O4 channel becomes highly disordered. This is accompanied by slight shifts of the nearby residues D1–D61 and CP43–E354 toward the position of W665, resulting in a narrowing of the space that has been occupied by W665 [Figs. 3(e) and 5(b)]. A water cluster (W546, W548, W612, W606 and W806) leading



**Table 2**

Interatomic distances of the OEC and their comparisons between different structures.

Values are averages of the nominal distances in two non-crystallographic symmetry related monomers.

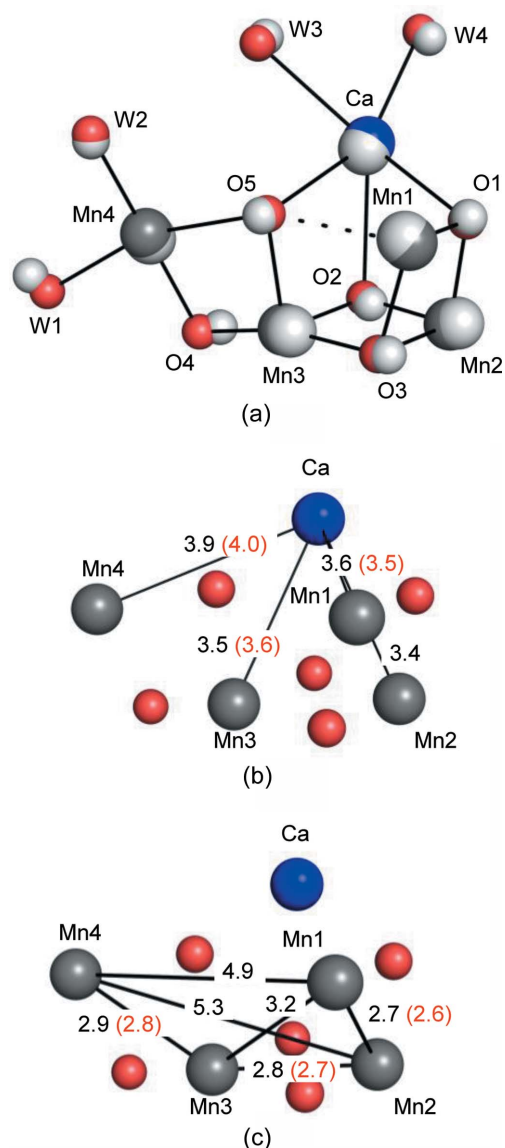
	The present study (room temperature)		100 K (Suga <i>et al.</i> , 2019)		Room temperature (Kern <i>et al.</i> , 2018)	
	S <sub>1</sub>	S <sub>2</sub>	S <sub>1</sub>	S <sub>2</sub>	S <sub>1</sub>	S <sub>2</sub>
Mn1–Mn2	2.66	2.64	2.60	2.68	2.78	2.81
Mn1–Mn3	3.16	3.15	3.16	3.21	3.25	3.26
Mn1–Mn4	4.94	4.90	4.97	4.90	4.86	4.86
Mn2–Mn3	2.76	2.72	2.72	2.75	2.85	2.84
Mn2–Mn4	5.28	5.26	5.27	5.20	5.21	5.24
Mn3–Mn4	2.86	2.84	2.89	2.76	2.74	2.74
Mn1–Ca	3.58	3.53	3.61	3.51	3.43	3.42
Mn2–Ca	3.44	3.44	3.42	3.40	3.38	3.41
Mn3–Ca	3.49	3.56	3.40	3.46	3.51	3.52
Mn4–Ca	3.90	4.04	3.76	3.90	3.83	3.90
Mn1–Mn2	2.66	2.64	2.60	2.68	2.78	2.81
Mn1–Mn3	3.16	3.15	3.16	3.21	3.25	3.26
Mn1–Mn4	4.94	4.90	4.97	4.90	4.86	4.86

to the luminal surface in the O4 channel also shifted its position during the S<sub>1</sub>-to-S<sub>2</sub> transition, and the shift of W548 induced structural changes of its hydrogen-bond partners D1-R334 and D1-N335 [Fig. 5(b)]. These changes in the O4 channel were similar to a previous study at cryogenic temperature (Suga *et al.*, 2019). However, another water molecule, W757, connected to W548 and W606 in this channel, was found to become disordered in the S<sub>1</sub>-to-S<sub>2</sub> transition at cryogenic temperature (Suga *et al.*, 2019), but this water molecule was not detected in either the S<sub>1</sub> or S<sub>2</sub> structures in the present study. Presumably, due to its peripheral location and thus its weaker association within the channel, W757 probably possesses higher mobility at room temperature, at which the TR-SFX experiments were conducted in the present study.

Another noticeable change observed near the OEC was a negative peak of  $-7.4\sigma$  covering W601, the hydrogen-bond donor to O1 and one of the members of a diamond-shaped water cluster in the O1 channel [Figs. 5(b) and 5(c)]. The negative peak of W601 has a slight displacement towards the O1 atom of OEC in the S<sub>2</sub>-state structure compared with its position in the dark structure. This indicates that W601 became disordered and shifted in the S<sub>1</sub>-to-S<sub>2</sub> transition. This change was not found in our previous study performed at cryogenic temperature (Suga *et al.*, 2019). Another SFX study at ambient temperature by Kern *et al.* (2018) reported shifts of three water molecules W601, W547 and W536 (W26, W27 and W30) of the diamond-shaped water cluster. In contrast, our previous fixed-target SFX study at cryogenic temperature reported that W571, a water molecule found at a cryogenic temperature only, became disordered instead of W601 (Suga *et al.*, 2019). Though we could not exclude the possibility of difference in the solution conditions under which the crystals were prepared and/or kept, or temperature employed (ambient versus cryogenic), these differences may indicate high mobility of the water molecules in the O1 channel during the S<sub>1</sub>-to-S<sub>2</sub> transition.

### 3.4. Q<sub>B</sub> and the non-heme iron site

After one-flash illumination, the second bound quinone electron acceptor Q<sub>B</sub> undergoes a reduction to form a stable plastosemiquinone intermediate Q<sub>B</sub><sup>•−</sup>. The isomorphous difference Fourier map showed a positive peak covering the Q<sub>B</sub> head and a pair of positive and negative peaks over the Q<sub>B</sub> tail, suggesting its slight movement during the S<sub>1</sub>-to-S<sub>2</sub> transition [Fig. 5(d)]. It should be noted that the peaks around the Q<sub>B</sub> site were smaller when compared with the OEC, or the O4 and O1 channels. Therefore, the motion of the residues discussed here can be confined to those showing a positional shift in coordinates and a pair of positive–negative difference



**Figure 6**  
(a)–(c) OEC structures in the S<sub>1</sub> and S<sub>2</sub> states are shown as gray and colored atoms, respectively. Color codes: blue for calcium; cyan for manganese; red for oxygen. The color codes for the OEC are the same in all figures unless otherwise noted. In (b) and (c), interatomic distances are given in Å, with the numbers in black for S<sub>1</sub> and red for S<sub>2</sub>. The distances in the S<sub>2</sub> state were shown only when they were greater by more than 0.1 Å compared with the corresponding distances in the S<sub>1</sub> state.

densities. The  $B$  factor of the  $Q_B$  head was decreased from  $123 \text{ \AA}^2$  in the  $S_1$  state to  $101 \text{ \AA}^2$  in the  $S_2$  state, and the features of the isomorphous difference Fourier map suggest a shortening of the hydrogen-bond interaction between the head of  $Q_B$  and D1-S264, indicating a tighter binding of the semiquinone  $Q_B^-$  to D1-S264. Thus, the first protonation of  $Q_B$  may occur at this site. This notion is consistent with the theoretical calculation that the first proton transfer from D1-S264 to  $Q_B$  is an energetically downhill process, whereas it is an uphill one from H215 to  $Q_B$  (Saito *et al.*, 2013). Several pairs of positive and negative peaks were also found around the hydrogen-bonded network formed by BCT, D1-Y246, D1-E244, D2-K264 and D2-E242, and BCT moved  $0.26 \text{ \AA}$  away from the non-heme iron [Fig. 5(d)]. These changes may be related to the reduction of the non-heme iron or proton uptake after one-flash illumination, which in turn suggests that these residues and BCT form part of the proton inlet channel for the protonation of  $Q_B^-$ .

#### 4. Discussion

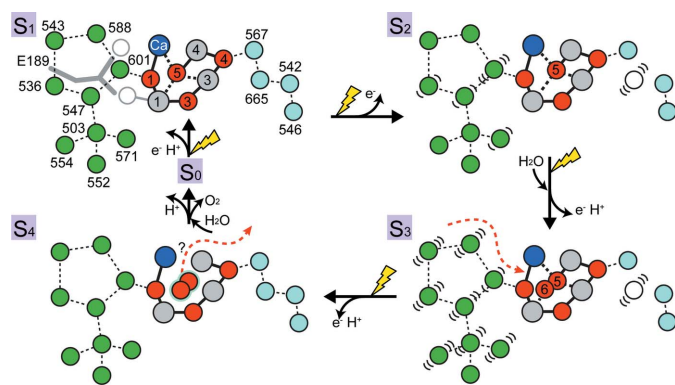
For the successful pump-probe TR-SFX experiments to analyze intermediate structures of proteins using light as the pump to initiate the reaction, it is critical to find an optimal light excitation condition with an optimal sample delivery system. These conditions vary considerably depending on the sample of interest as well as the setup of the pump-probe SFX system. Under optimal conditions, one can save a large amount of valuable protein sample as well as XFEL beam time for obtaining structures of the targeted intermediate states. In the present study, we showed a method to determine a boundary excited by the pump by altering the flow rate of the microcrystals and evaluating the resulting light-induced structural changes of PSII. It was shown that the optical illumination area of the laser is scattered into an area well beyond that defined by the laser spot itself, and the flow rate of the crystal stream has to be fast enough to avoid light contamination by the preceding flash illumination. This calls for caution in defining the optical illumination conditions in TR-SFX experiments (Nogly *et al.*, 2018; Kovacs *et al.*, 2019; Grünbein *et al.*, 2020). With the optimal conditions obtained, we analyzed the structure of PSII in the  $S_2$  state at room temperature by the TR-SFX method. This method can be applied to other samples, whose catalytic reaction is triggered by light, using systems which deliver microcrystals in a continuous flow. Importantly, this method also provides suitable light conditions for samples, such as PSII, that require successive flash excitations for the catalytic cycle.

The PSII structures in the  $S_1$  and  $S_2$  states, as well as the isomorphous difference Fourier map obtained in the present study, revealed the light-induced structural changes localized at the OEC,  $Q_B$  site and water molecules in the O1 and O4 channels. Comparison of the changes observed at ambient temperature with  $\Delta t = 10 \text{ ms}$  obtained in the present study, with those of the previous study, where the  $S_2$  state was trapped cryogenically with  $\Delta t \simeq 1 \text{ s}$  and the diffraction data collected at low temperature (Suga *et al.*, 2019), as well as that

carried out at ambient temperature with  $\Delta t = 200 \text{ ms}$  (Kern *et al.*, 2018), shows an overall tendency of the structural changes. The changes found in the OEC were identical regardless of the temperature or the different  $\Delta t$ , suggesting the stability and long lifetime of the catalytic center in the  $S_2$  state. By contrast, the changes found in some amino acid residues and water molecules differ substantially either depending on the temperatures or  $\Delta t$ , reflecting a rather large mobility of the protein environment. This tendency is apparent in the  $Q_B$  site, in which the quinone changes its redox forms during the catalytic cycle. The changes in the  $Q_B$  site were larger with  $\Delta t = 10 \text{ ms}$  but smaller with  $\Delta t = 200 \text{ ms}$  (Kern *et al.*, 2018) and further diminished at  $\Delta t \simeq 1 \text{ s}$  (Suga *et al.*, 2019). However, the  $Q_B$  yield based on the oxidation status of the non-heme iron may be different under experimental conditions.

Comparisons of the structural changes in the O1 and O4 channels with previous crystallographic studies (Suga *et al.*, 2019; Kern *et al.*, 2018) also reveal the differences in the channels as a function of the  $S_i$ -state. W665, the second water from O4 in the O4 channel, became highly mobile in the  $S_1$ -to- $S_2$  transition, which is consistent with previous studies. The disorder of W665 breaks the hydrogen-bonded network in the O4 channel in the  $S_2$  state. In our initial report of the disorder of W665 in the  $S_2$  state, we proposed that the proton transfer may occur in the O4 channel in the  $S_1$ -to- $S_2$  transition (Suga *et al.*, 2019). The changes observed in the O4 channel were similar to the theoretical calculation that reported a proton release in the channel in the  $S_0$ -to- $S_1$  transition (Saito *et al.*, 2015). However, FTIR studies on the D1-S169A mutant, which perturbs the interaction of a water molecule hydrogen bonded to O4, showed only minor effects on the efficiencies or kinetics of the  $S_i$ -state transitions (Shimada *et al.*, 2020; Ghosh *et al.*, 2019). In contrast, the  $S_2$  state multiline EPR spectrum of the D1-S169A mutant differs significantly from that of wild-type (Ghosh *et al.*, 2019). A further theoretical study on the O4 channel showed that removing water molecules in the O4 channel results in a decrease in the  $S_2/S_1$  redox potential by  $\sim 80 \text{ mV}$  (Mandal *et al.*, 2020). Therefore, the weakened hydrogen-bonded network resulting from the increased mobility of W665 may be necessary for the decrease in the redox potential to facilitate the  $S_1$ -to- $S_2$  transition.

In contrast to the changes observed in the O4 channel, the changes that occurred in the water molecules of the O1 channel differ between the present and previous studies. W601, the hydrogen-bond donor to O1, became disordered in the present study at  $\Delta t = 10 \text{ ms}$ , and two additional water molecules W547 and W536 (W26, W27 and W29) (Kern *et al.*, 2018) became disordered at  $\Delta t = 200 \text{ ms}$ . However, instead of these water molecules, W571 became disordered in the study with  $\Delta t \simeq 1 \text{ s}$  at cryogenic temperature (Suga *et al.*, 2019). Because the O1 channel is a broad channel where water molecules are the major component of the hydrogen-bonded network, substrate water molecules are likely to propagate through the water molecules in the O1 channel. The differences observed in the structural changes of water molecules located in this channel at different temperatures and/or delay times after laser flash illumination may reflect the high



**Figure 7**

Structural changes in the OEC and water molecules found in the present and previous studies are summarized. Red and gray spheres with numbers in the OEC are oxygen and manganese atoms, respectively, and the Mn2 and O2 atoms have been omitted for clarity.

mobility of water molecules in this channel. Here, we could not exclude the possibility that the structural differences were caused by the difference in the solution conditions in which the crystals were made and/or kept. It should be noted that glycerol or DMSO molecules used as cryo-protectants were indeed identified in the O1 channel (Umena *et al.*, 2011; Suga *et al.*, 2015), suggesting the entry of outside molecules into this channel. Thus, increasing the mobility of water in this channel seems beneficial for driving the incorporation of substrate water into the catalytic site in the following S<sub>3</sub> state.

Fig. 7 summarizes the structural changes regarding the water molecules in the O1 and O4 channels observed in the present study. In the S<sub>1</sub>-to-S<sub>2</sub> transition, W665 in the O4 channel becomes highly mobile, disconnecting the 15 Å-long water chain starting from O4-W567. W665 remains disordered in the S<sub>3</sub> state but becomes re-ordered beyond the S<sub>3</sub> state (Suga *et al.*, 2019; Kern *et al.*, 2018). W601 in the O1 channel also becomes disordered in the S<sub>1</sub>-to-S<sub>2</sub> transition. The mobility of W601 is propagated to W547, W536 and W571 through the hydrogen-bonded network in the O1 channel. D1-E189, the only monodentate carboxyl ligand of the OEC, flips by 0.5 Å in the following S<sub>2</sub>-to-S<sub>3</sub> transition, leading to broadening of the O1 channel, thereby further increasing the mobility of water molecules there (Suga *et al.*, 2017, 2019; Ibrahim *et al.*, 2020). The flipping of D1-E189 makes a space in the vicinity of O5 to allow the insertion of the additional water (oxygen) O6 in the subsequent S<sub>2</sub>-to-S<sub>3</sub> transition. This suggests that water molecules in the O1 channel may be the source for O6. Dioxygen is then probably formed between O5 and O6 by an oxyl/oxo coupling mechanism, or between O5 and another water molecule, and in this case, O6 refills the empty substrate site. The changes of water molecules in the O1 channel support the mechanism of water delivery from the Ca side (Kim & Debus, 2019; Sakamoto *et al.*, 2017; Suga *et al.*, 2019), rather than the O4 channel (Kawashima *et al.*, 2018; Wang *et al.*, 2017; Retegan *et al.*, 2016). Further structural analysis will be required to reveal how the OEC incorporates those water molecules in the O1 channel and how protons egress during the subsequent S-state transitions.

## Acknowledgements

We thank T. Nakano and S. Goda for growing and harvesting the cyanobacterial cells. The XFEL experiments were performed at beamlines 2 and 3 of SACLA with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (proposals 2016A8033, 2016B8066, 2017A8041, 2017B8025, 2018A8037 and 2018B8029) and we thank the staff at SACLA for their help. We also acknowledge the computational support from SACLA HPC system and Mini-K supercomputer system.

## Funding information

The following funding is acknowledged: Japan Society for the Promotion of Science (grant Nos. JP17H05884, JP20H05446, JP20H03226 to MS; grant No. JP17H06434 to JRS; grant No. JP19H05784 to MK; grant No. JP19H05781 to EN); Japan Agency for Medical Research and Development (grant awarded to SI); Japan Science and Technology Agency, Precursory Research for Embryonic Science and Technology (grant No. JPMJPR18G8 to MS; grant No. JPMJPR16P1 to FA); Platform Project for Supporting Drug Discovery and Life Science Research [Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)] from AMED (grant No. JP20am0101070).

## References

- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* **D66**, 213–221.
- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* **D68**, 352–367.
- Barends, T. R., Foucar, L., Ardevol, A., Nass, K., Aquila, A., Botha, S., Doak, R. B., Falahati, K., Hartmann, E., Hilpert, M., Heinz, M., Hoffmann, M. C., Kofinger, J., Koglin, J. E., Kovacsova, G., Liang, M., Milathianaki, D., Lemke, H. T., Reinstein, J., Roome, C. M., Shoeman, R. L., Williams, G. J., Burghardt, I., Hummer, G., Boutet, S. & Schlichting, I. (2015). *Science*, **350**, 445–450.
- Barty, A., Kirian, R. A., Maia, F. R. N. C., Hantke, M., Yoon, C. H., White, T. A. & Chapman, H. (2014). *J. Appl. Cryst.* **47**, 1118–1131.
- Brewster, A. S., Waterman, D. G., Parkhurst, J. M., Gildea, R. J., Young, I. D., O’Riordan, L. J., Yano, J. & Sauter, N. K. (2018). *Acta Cryst.* **D74**, 877–894.
- Collaborative Computational Project, Number 4, (1994). *Acta Cryst.* **D50**, 760–763.
- Cox, N., Pantazis, D. A. & Lubitz, W. (2020). *Annu. Rev. Biochem.* **89**, 795–820.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Ghosh, I., Banerjee, G., Kim, C. J., Reiss, K., Batista, V. S., Debus, R. J. & Brudvig, G. W. (2019). *Biochemistry*, **58**, 1379–1387.
- Grünbein, M. L., Stricker, M., Nass Kovacs, G., Kloos, M., Doak, R. B., Shoeman, R. L., Reinstein, J., Lecler, S., Haacke, S. & Schlichting, I. (2020). *Nat. Methods*, **17**, 681–684.
- Ibrahim, M., Fransson, T., Chatterjee, R., Cheah, M. H., Hussein, R., Lassalle, L., Sutherlin, K. D., Young, I. D., Fuller, F. D., Gul, S., Kim, I. S., Simon, P. S., de Lichtenberg, C., Chernev, P., Bogacz, I., Pham, C. C., Orville, A. M., Saichek, N., Northen, T., Batyuk, A., Carbajo, S., Alonso-Mori, R., Tono, K., Owada, S., Bhowmick, A., Bolotovskiy, R., Mendez, D., Moriarty, N. W., Holton, J. M.,



- Dobbek, H., Brewster, A. S., Adams, P. D., Sauter, N. K., Bergmann, U., Zouni, A., Messinger, J., Kern, J., Yachandra, V. K. & Yano, J. (2020). *Proc. Natl Acad. Sci. USA*, **117**, 12624–12635.
- Kato, Y., Akita, F., Nakajima, Y., Suga, M., Umena, Y., Shen, J. R. & Noguchi, T. (2018). *J. Phys. Chem. Lett.* **9**, 2121–2126.
- Kawashima, K., Takaoka, T., Kimura, H., Saito, K. & Ishikita, H. (2018). *Nat. Commun.* **9**, 1247.
- Kern, J., Alonso-Mori, R., Tran, R., Hattne, J., Gildea, R. J., Echols, N., Glöckner, C., Hellmich, J., Laksmono, H., Sierra, R. G., Lassalle-Kaiser, B., Koroidov, S., Lampe, A., Han, G., Gul, S., DiFiore, D., Milathianaki, D., Fry, A. R., Miahnahri, A., Schafer, D. W., Messerschmidt, M., Seibert, M. M., Koglin, J. E., Sokaras, D., Weng, T. C., Sellberg, J., Latimer, M. J., Grosse-Kunstleve, R. W., Zwart, P. H., White, W. E., Glatzel, P., Adams, P. D., Bogan, M. J., Williams, G. J., Boutet, S., Messinger, J., Zouni, A., Sauter, N. K., Yachandra, V. K., Bergmann, U. & Yano, J. (2013). *Science*, **340**, 491–495.
- Kern, J., Chatterjee, R., Young, I. D., Fuller, F. D., Lassalle, L., Ibrahim, M., Gul, S., Fransson, T., Brewster, A. S., Alonso-Mori, R., Hussein, R., Zhang, M., Douthit, L., de Lichtenberg, C., Cheah, M. H., Shevela, D., Wersig, J., Seuffert, I., Sokaras, D., Pastor, E., Weninger, C., Kroll, T., Sierra, R. G., Aller, P., Butryn, A., Orville, A. M., Liang, M., Batyuk, A., Koglin, J. E., Carbajo, S., Boutet, S., Moriarty, N. W., Holton, J. M., Dobbek, H., Adams, P. D., Bergmann, U., Sauter, N. K., Zouni, A., Messinger, J., Yano, J. & Yachandra, V. K. (2018). *Nature*, **563**, 421–425.
- Kern, J., Tran, R., Alonso-Mori, R., Koroidov, S., Echols, N., Hattne, J., Ibrahim, M., Gul, S., Laksmono, H., Sierra, R. G., Gildea, R. J., Han, G., Hellmich, J., Lassalle-Kaiser, B., Chatterjee, R., Brewster, A. S., Stan, C. A., Glöckner, C., Lampe, A., DiFiore, D., Milathianaki, D., Fry, A. R., Seibert, M. M., Koglin, J. E., Gallo, E., Uhlrig, J., Sokaras, D., Weng, T. C., Zwart, P. H., Skinner, D. E., Bogan, M. J., Messerschmidt, M., Glatzel, P., Williams, G. J., Boutet, S., Adams, P. D., Zouni, A., Messinger, J., Sauter, N. K., Bergmann, U., Yano, J. & Yachandra, V. K. (2014). *Nat. Commun.* **5**, 4371.
- Kim, C. J. & Debus, R. J. (2019). *Biochemistry*, **58**, 3185–3192.
- Kok, B. F. B., Forbush, B. & McGloin, M. (1970). *Photochem. Photobiol.* **11**, 457–475.
- Kubo, M., Nango, E., Tono, K., Kimura, T., Owada, S., Song, C., Mafuné, F., Miyajima, K., Takeda, Y., Kohno, J., Miyauchi, N., Nakane, T., Tanaka, T., Nomura, T., Davidsson, J., Tanaka, R., Murata, M., Kameshima, T., Hatsui, T., Joti, Y., Neutze, R., Yabashi, M. & Iwata, S. (2017). *J. Synchrotron Rad.* **24**, 1086–1091.
- Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N. A., Rendek, K. N., Hunter, M. S., Shoeman, R. L., White, T. A., Wang, D., James, D., Yang, J.-H., Cobb, D. E., Reeder, B., Sierra, R. G., Liu, H., Barty, A., Aquila, A. L., Deponte, D., Kirian, R. A., Bari, S., Bergkamp, J. J., Beyerlein, K. R., Bogan, M. J., Caleman, C., Chao, T.-C., Conrad, C. E., Davis, K. M., Fleckenstein, H., Galli, L., Hau-Riege, S. P., Kassemeyer, S., Laksmono, H., Liang, M., Lomb, L., Marchesini, S., Martin, A. V., Messerschmidt, M., Milathianaki, D., Nass, K., Ros, A., Roy-Chowdhury, S., Schmidt, K., Seibert, M., Steinbrener, J., Stellato, F., Yan, L., Yoon, C., Moore, T. A., Moore, A. L., Pushkar, Y., Williams, G. J., Boutet, S., Doak, R. B., Weierstall, U., Frank, M., Chapman, H. N., Spence, J. C. H. & Fromme, P. (2014). *Nature*, **513**, 261–265.
- Mandal, M., Kawashima, K., Saito, K. & Ishikita, H. (2020). *J. Phys. Chem. Lett.* **11**, 249–255.
- Nakane, T., Joti, Y., Tono, K., Yabashi, M., Nango, E., Iwata, S., Ishitani, R. & Nureki, O. (2016). *J. Appl. Cryst.* **49**, 1035–1041.
- Nango, E., Kubo, M., Tono, K. & Iwata, S. (2019). *Appl. Sci.* **9**, 5505.
- Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., Shimamura, T., Im, D., Fujiwara, T., Yamanaka, Y., Jeon, B., Nishizawa, T., Oda, K., Fukuda, M., Andersson, R., Båth, P., Dods, R., Davidsson, J., Matsuoka, S., Kawatake, S., Murata, M., Nureki, O., Owada, S., Kameshima, T., Hatsui, T., Joti, Y., Schertler, G., Yabashi, M., Bondar, A. N., Standfuss, J., Neutze, R. & Iwata, S. (2016). *Science*, **354**, 1552–1557.
- Nass Kovacs, G., Colletier, J.-P., Grünbein, M. L., Yang, Y., Stensitzki, T., Batyuk, A., Carbajo, S., Doak, R. B., Ehrenberg, D., Foucar, L., Gasper, R., Gorel, A., Hilpert, M., Kloos, M., Koglin, J. E., Reinstein, J., Roome, C. M., Schlesinger, R., Seaberg, M., Shoeman, R. L., Stricker, M., Boutet, S., Haacke, S., Heberle, J., Heyne, K., Domratcheva, T., Barends, T. R. M. & Schlichting, I. (2019). *Nat. Commun.* **10**, 3177.
- Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature*, **406**, 752–757.
- Nogly, P., Weinert, T., James, D., Carbajo, S., Ozerov, D., Furrer, A., Gashi, D., Borin, V., Skopintsev, P., Jaeger, K., Nass, K., Bath, P., Bosman, R., Koglin, J., Seaberg, M., Lane, T., Kekilli, D., Brunle, S., Tanaka, T., Wu, W., Milne, C., White, T., Barty, A., Weierstall, U., Panneels, V., Nango, E., Iwata, S., Hunter, M., Schapiro, I., Schertler, G., Neutze, R. & Standfuss, J. (2018). *Science*, **361**, eaat0094.
- Pande, K., Hutchison, C. D., Groenhof, G., Aquila, A., Robinson, J. S., Tenboer, J., Basu, S., Boutet, S., DePonte, D. P., Liang, M., White, T. A., Zatsepin, N. A., Yefanov, O., Morozov, D., Oberthuer, D., Gati, C., Subramanian, G., James, D., Zhao, Y., Koralek, J., Brayshaw, J., Kupitz, C., Conrad, C., Roy-Chowdhury, S., Coe, J. D., Metz, M., Xavier, P. L., Grant, T. D., Koglin, J. E., Ketawala, G., Fromme, R., rajer, V., Henning, R., Spence, J. C., Ourmazd, A., Schwander, P., Weierstall, U., Frank, M., Fromme, P., Barty, A., Chapman, H. N., Moffat, K., van Thor, J. J. & Schmidt, M. (2016). *Science*, **352**, 725–729.
- Retegan, M., Krewald, V., Mamedov, F., Neese, F., Lubitz, W., Cox, N. & Pantazis, D. A. (2016). *Chem. Sci.* **7**, 72–84.
- Saito, K., Rutherford, A. W. & Ishikita, H. (2013). *Proc. Natl Acad. Sci. USA*, **110**, 954–959.
- Saito, K., William Rutherford, A. & Ishikita, H. (2015). *Nat. Commun.* **6**, 8488.
- Sakamoto, H., Shimizu, T., Nagao, R. & Noguchi, T. (2017). *J. Am. Chem. Soc.* **139**, 2022–2029.
- Sauter, N. K. (2015). *J. Synchrotron Rad.* **22**, 239–248.
- Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. (2013). *Acta Cryst. D* **69**, 1274–1282.
- Shen, J.-R. (2015). *Annu. Rev. Plant Biol.* **66**, 23–48.
- Shen, J.-R. & Kamiya, N. (2000). *Biochemistry*, **39**, 14739–14744.
- Shimada, Y., Kitajima-Ihara, T., Nagao, R. & Noguchi, T. (2020). *J. Phys. Chem. B*, **124**, 1470–1480.
- Shimazu, Y., Tono, K., Tanaka, T., Yamanaka, Y., Nakane, T., Mori, C., Terakado Kimura, K., Fujiwara, T., Sugahara, M., Tanaka, R., Doak, R. B., Shimamura, T., Iwata, S., Nango, E. & Yabashi, M. (2019). *J. Appl. Cryst.* **52**, 1280–1288.
- Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H. & Shen, J.-R. (2015). *Nature*, **517**, 99–103.
- Suga, M., Akita, F., Sugahara, M., Kubo, M., Nakajima, Y., Nakane, T., Yamashita, K., Umena, Y., Nakabayashi, M., Yamane, T., Nakano, T., Suzuki, M., Masuda, T., Inoue, S., Kimura, T., Nomura, T., Yonekura, S., Yu, L. J., Sakamoto, T., Motomura, T., Chen, J. H., Kato, Y., Noguchi, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Nango, E., Tanaka, R., Naitow, H., Matsuura, Y., Yamashita, A., Yamamoto, M., Nureki, O., Yabashi, M., Ishikawa, T., Iwata, S. & Shen, J. R. (2017). *Nature*, **543**, 131–135.
- Suga, M., Akita, F., Yamashita, K., Nakajima, Y., Ueno, G., Li, H., Yamane, T., Hirata, K., Umena, Y., Yonekura, S., Yu, L. J., Murakami, H., Nomura, T., Kimura, T., Kubo, M., Baba, S., Kumasaka, T., Tono, K., Yabashi, M., Isobe, H., Yamaguchi, K., Yamamoto, M., Ago, H. & Shen, J. R. (2019). *Science*, **366**, 334–338.
- Suga, M., Shimada, A., Akita, F., Shen, J. R., Toshi, T. & Sugimoto, H. (2020). *Biochim. Biophys. Acta Gen. Subj.* **1864**, 129466.
- Sugahara, M., Mizohata, E., Nango, E., Suzuki, M., Tanaka, T., Masuda, T., Tanaka, R., Shimamura, T., Tanaka, Y., Suno, C., Ihara,

- K., Pan, D., Kakinouchi, K., Sugiyama, S., Murata, M., Inoue, T., Tono, K., Song, C., Park, J., Kameshima, T., Hatsui, T., Joti, Y., Yabashi, M. & Iwata, S. (2015). *Nat. Methods*, **12**, 61–63.
- Tenboer, J., Basu, S., Zatsepin, N., Pande, K., Milathianaki, D., Frank, M., Hunter, M., Boutet, S., Williams, G. J., Koglin, J. E., Oberthuer, D., Heymann, M., Kupitz, C., Conrad, C., Coe, J., Roy-Chowdhury, S., Weierstall, U., James, D., Wang, D., Grant, T., Barty, A., Yefanov, O., Scales, J., Gati, C., Seuring, C., Srajer, V., Henning, R., Schwander, P., Fromme, R., Ourmazd, A., Moffat, K., Van Thor, J. J., Spence, J. C., Fromme, P., Chapman, H. N. & Schmidt, M. (2014). *Science*, **346**, 1242–1246.
- Tono, K., Nango, E., Sugahara, M., Song, C., Park, J., Tanaka, T., Tanaka, R., Joti, Y., Kameshima, T., Ono, S., Hatsui, T., Mizohata, E., Suzuki, M., Shimamura, T., Tanaka, Y., Iwata, S. & Yabashi, M. (2015). *J. Synchrotron Rad.* **22**, 532–537.
- Umena, Y., Kawakami, K., Shen, J. R. & Kamiya, N. (2011). *Nature*, **473**, 55–60.
- Wang, J., Askerka, M., Brudvig, G. W. & Batista, V. S. (2017). *ACS Energy Lett.* **8**, 2299–2306.
- Waterman, D. G., Winter, G., Gildea, R. J., Parkhurst, J. M., Brewster, A. S., Sauter, N. K. & Evans, G. (2016). *Acta Cryst.* **D72**, 558–575.
- White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A. & Chapman, H. N. (2012). *J. Appl. Cryst.* **45**, 335–341.
- Young, I. D., Ibrahim, M., Chatterjee, R., Gul, S., Fuller, F. D., Koroidov, S., Brewster, A. S., Tran, R., Alonso-Mori, R., Kroll, T., Michels-Clark, T., Laksmono, H., Sierra, R. G., Stan, C. A., Hussein, R., Zhang, M., Douthit, L., Kubin, M., de Lichtenberg, C., Vo Pham, L., Nilsson, H., Cheah, M. H., Shevela, D., Saracini, C., Bean, M. A., Seuffert, I., Sokaras, D., Weng, T. C., Pastor, E., Weninger, C., Fransson, T., Lassalle, L., Bräuer, P., Aller, P., Docker, P. T., Andi, B., Orville, A. M., Glowina, J. M., Nelson, S., Sikorski, M., Zhu, D., Hunter, M. S., Lane, T. J., Aquila, A., Koglin, J. E., Robinson, J., Liang, M., Boutet, S., Lyubimov, A. Y., Uervirojnangkoorn, M., Moriarty, N. W., Liebschner, D., Afonine, P. V., Waterman, D. G., Evans, G., Wernet, P., Dobbek, H., Weis, W. I., Brunger, A. T., Zwart, P. H., Adams, P. D., Zouni, A., Messinger, J., Bergmann, U., Sauter, N. K., Kern, J., Yachandra, V. K. & Yano, J. (2016). *Nature*, **540**, 453–457.