

Heterogeneous Composition of Oxygen-Evolving Complexes in Crystal Structures of Dark-Adapted Photosystem II

Jimin Wang,* Christopher J. Gisriel, Krystle Reiss, Hao-Li Huang, William H. Armstrong, Gary W. Brudvig, and Victor S. Batista



Cite This: *Biochemistry* 2021, 60, 3374–3384



Read Online

ACCESS |



Metrics & More

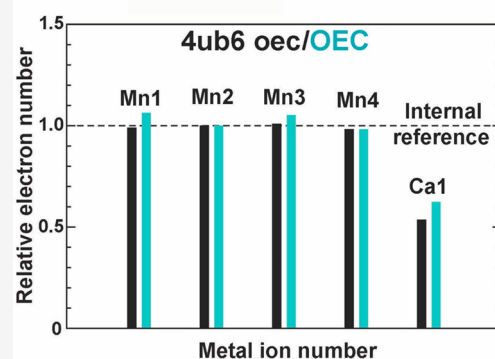


Article Recommendations



Supporting Information

ABSTRACT: Photosystem II (PSII) is a homodimeric protein complex that catalyzes water oxidation at the oxygen-evolving complex (OEC), a heterocubanoide calcium–tetramanganese cluster. Here, we analyze the omit electron density peaks of the OEC's metal ions in five X-ray free-electron laser PSII structures at resolutions between 2.15 and 1.95 Å. The omit peaks can be described by the total number of electrons and approximated by the variance of electron density distribution when the distributions are spherically symmetric. We show that the number of electrons of metal centers is different in the two OECs of PSII dimers, implying that the oxidation states and/or occupancies of individual metal ions are different in the two monomers. In either case, we find that the two OECs of dark-adapted PSII dimers in crystals are not fully synchronized in the S_1 state. Differences in redox states of the OEC in PSII only partially account for the observation that the electron densities integrate to a smaller number of electrons than expected. Differences between the determined and expected relative electron numbers are much larger than the estimated errors, indicating heterogeneity in the OEC composition.



The oxygen-evolving complex (OEC) of photosystem II (PSII) splits two water molecules during each turn of its catalytic cycle, producing one dioxygen (O_2) molecule, four protons, and four electrons.^{1–4} In the classic Joliot–Kok model, the OEC accumulates four oxidizing equivalents, advancing its redox state from S_0 to S_4 .⁴ Oxygen is then rapidly released and the OEC returns to the S_0 state, completing the cycle. While the S_4 intermediate is so transient that it has yet to be detected, the S_0 – S_3 states of the OEC have all been characterized using a variety of biophysical and biochemical methods.^{5–7} In particular, the S_2 and S_3 states of the OEC are unstable *in vitro* and return to the S_1 state within minutes in the dark.⁴

The X-ray diffraction (XRD) model of the dark-adapted PSII dimers from *Thermosynechococcus vulcanus* at 1.90 Å resolution revealed that the OEC structure contains a Mn_4CaO_5 cluster that has oxygen ligands bridging metal centers.⁸ Density functional theory quantum mechanics/molecular mechanics (QM/MM) analysis suggested that the composition of the XRD model corresponded to a mixture of oxidation states, including species more reduced than those observed in the catalytic cycle of water splitting in addition to a component of S_1 and ~60% of the S_0 intermediate.⁹ Later, Kamiya and colleagues analyzed two crystal structures of dark-adapted PSII dimers using a low-dose synchrotron X-ray source at 1.85 and 1.87 Å resolution,¹⁰ and Shen and colleagues obtained two structures using X-ray free-electron laser (XFEL) crystallography at 1.95 Å resolution.¹¹ The analysis showed that the two OEC monomers of the crystallographic asymmetric unit exhibited substantially differ-

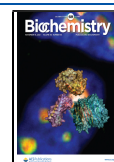
ent geometries (in Kamiya's and Shen's high-resolution PSII structures).¹⁰ However, the origin of structural differences between the two monomers remains unknown. Consistent with these observations, Ca ions in OECs of some XFEL structures exhibited *B*-factors unusually higher than those of the Mn ions,¹¹ suggesting lower occupancy because the *B*-factors of bonded atoms in those structures should not vary as much.

High-resolution electron density (ED) maps obtained by X-ray crystallography contain quantitative information about both the number of electrons per atom (EN) and the variance of the ED distribution when the ED peaks of individual scattering centers become resolved, particularly for metal centers with high atomic numbers. Recently, we developed an analytical method for estimating the ENs associated with peaks in crystallographic σ_A -weighted $2F_o - F_c$ ED maps of proteins.^{12–14} Here, we implement a variation of that method to analyze the σ_A -weighted $F_o - F_c$ metal ion-omit maps in two steps. First, we validate the method on a benchmark system of fully occupied metal ions from nitrogenase crystal structures.^{15–17} Second, we apply the method to analyze dark-adapted cyanobacterial PSII structures

Received: September 13, 2021

Revised: October 16, 2021

Published: October 29, 2021



obtained using XFEL crystallography.^{11,18–20} We find evidence that the two OECs of dark-adapted PSII crystals have, on average, different metal ion compositions and are in different oxidation states. Therefore, we conclude that a proper interpretation of both dark-adapted structures and light-induced reaction intermediates generated from dark-adapted crystals must take into account the fact that a significant fraction of the OECs may not be in the expected state, as has often been assumed to be the case.

THEORY, MATERIALS, AND ANALYTIC PROCEDURES

Crystallographic Theory on Electron Density Peaks.

ED peaks for nearly all non-H atoms are resolved in well-refined large-protein crystal structures such as *Escherichia coli* catalase at ~ 1.7 Å resolution.¹² Single-electron differences among O, N, and C atoms of individual residues can be directly recognized by the size of their corresponding ED peaks, which is often a spherical envelope at a given contour level.¹² ED peaks for electron-dense metal ion ions in protein structures at ~ 2.0 Å resolution are also well resolved, particularly in metal ion-omit $F_o - F_c$ ED maps omitting one metal ion atom of interest at a time (see below).¹² These peaks are often spherically symmetric, and the logarithm of the density value as a function of distance squared (r^2) to the peak center (i.e., $\log rr$ plot) is largely a linear function. Hence, the ED distribution within the peaks can be approximated as a single Gaussian function at ~ 2.0 Å resolution. These peaks contain two quantitative parameters: the variance of the ED distribution and the total EN, both of which can be fitted analytically relative to an internal reference as described elsewhere.^{12–14} With increasing resolution, multiple Gaussian functions are needed to describe the ED peaks.^{21–23} When the query and reference have the same multiple Gaussian functions (for example, redox inactive Mg^{2+} or Mn^{2+} bound in DNA polymerases¹³), their $\log rr$ plots will have exactly the same features. When their initial slopes are fitted, the entire curves will also be fitted.¹³ In this case, the same analytic procedure can be used to determine relative occupancies accurately if occupancies are variables.¹³

For spherically symmetric ED peaks (ρ), the analytic expression of a multiple-Gaussian ED function in a polar coordinate system is as follows.^{12,24}

$$\rho(r, B) = 8\pi^{3/2} \sum_{i=1}^s a_i e^{-4\pi^2 r^2 / (B+b_i)} / (B + b_i)^{3/2} \quad (1)$$

$$\rho(0, B) = 8\pi^{3/2} \sum_{i=1}^s a_i / (B + b_i)^{3/2} \quad (2)$$

where (a_i, b_i) are Gaussian function parameters ($i = 1, 2, 3, 4$, and S , and $b_5 = 0$) and B is an atomic displacement parameter. These Gaussian function parameters are the same as those of X-ray form factors (f) of the atom in question in the absence or presence of a B -factor:^{25–27}

$$f(s) = \sum_{i=1}^s a_i e^{-b_i s^2} \quad (3)$$

$$f(s, B) = f(s) e^{-Bs^2} = \sum_{i=1}^s a_i e^{-(B+b_i)s^2} \quad (4)$$

$$f(0) = \sum_{i=1}^s a_i \quad (5)$$

where s is the scattering vector $\sin \theta / \lambda$. The total EN is the forward scattering length at $\sin \theta / \lambda = 0$, and it is also the sum of all Gaussian function linear coefficients.

Unfortunately, an analytic expression of form factors is not available for the Mn^{4+} ions that are present in the intermediate states of the OEC of PSII.^{9,28} We solve this problem by using Mn omit $F_o - F_c$ ED maps in which the unknown parameters of its form factors are no longer needed to evaluate the ED function or its B -factor.

With increasing atomic B -factor ($B = 8\pi^2 \langle |\Delta r|^2 \rangle$), the ED peak spreads outward; B contributes to the variance of the ED distribution, and b_i describes the intrinsic variance. When $B \gg b_i$, $B + b_i \approx B$, and multiple Gaussian terms gradually degenerate into a single term in both reciprocal space and real space. In the case of PSII at ~ 2.0 Å resolution, approximately 500,000 reflections contribute to each Mn omit ED peak.⁸ Even so, there is insufficient resolution to resolve individual Gaussian terms. We can determine the overall variance only from the mean Gaussian terms ($\langle B + b_i \rangle$). Analytically fitted Gaussian parameters cannot distinguish between b_i and B terms, nor can we decompose them unless diffraction data are measured at sub-angstrom resolution with the atomic B value becoming very small.^{22,23,29}

In the absence of model phase errors (as in error-free computer simulations), the maximal size of the omit $F_o - F_c$ ED peak is only approximately half of its full ED peak,³⁰ which can be evaluated using vector difference Fourier maps that would require knowledge of form factors and the B -factor of the atom in question. In the presence of modeling errors, the size of the ED peak decreases further. Therefore, the analytic method described here is not designed for determination of the absolute ENs with respect to the theoretical ED peak. It is suitable only for determination of relative ENs to an internal reference. Modeling errors away from the OEC affect both the reference and the query (or unknown) the same way and hence are largely canceled out.

Successful determination of relative ENs due to varying occupancies of redox inactive Mn^{2+} ions bound in DNA polymerases and for chemical identification of an atom in a modified protein residue based on σ_A -weighted $2F_o - F_c$ ED maps has been demonstrated elsewhere.^{13,14} Because this is the first time that we are applying our analytic method to σ_A -weighted $F_o - F_c$ ED peaks, we decided to extensively calibrate the new method using nitrogenase before applying it to PSII.

Materials and Analytic Procedures. Nitrogenase has an $\alpha_2\beta_2$ subunit structure. At the interface of each $\alpha\beta$ pair resides a P cluster that has a composition of Fe_8S_7 , and each α subunit contains a so-called M cluster ($FeMo$ cofactor) consisting of one molybdenum atom, one carbon atom, seven iron atoms, and nine sulfur atoms (Figure S1).¹⁵ Five nitrogenase X-ray structures and their corresponding diffraction data sets were retrieved from the Protein Data Bank (PDB): entries 3u7q at 1.00 Å resolution, 4wes at 1.08 Å resolution, and 1qgu, 1qh8, and 1qh1 at 1.60 Å resolution.^{15–17} Metal ion-omit difference Fourier maps were calculated using Refmac5 in the CCP4 suite by setting both the refinement cycle and the occupancy of the selected metal ion to zero.^{31,32} The resulting maps were visualized, and the peak centers of the omitted metal ions were determined using the first moment with Coot.³³

Five XFEL PSII models and crystallographic data sets were retrieved from PDB³⁴ entries 6jly at 2.15 Å resolution, 6dhe at 2.05 Å resolution, 4ub6 at 1.95 Å resolution, 4ub8 at 1.95 Å resolution, and 6w1o at 2.08 Å resolution.^{11,18–20} For PSII structures, alignment of the two OECs of each PSII dimer was carried out by least-squares superposition of the ED using the Rave package and Chimera.^{35,36}

Spherical averaging of ED was carried out using numerical density averaging from 12 sampling points arranged with cuboctahedral symmetry on a sphere of increasing radius and plotted as a function of radius in an increment of 0.01 Å.¹³ The deviation from spherical symmetry was measured using ± 1 standard deviation from the mean value and examined to ensure the validity of using spherical symmetry. The ED function is normalized as a percentage of its peak maximum, $\rho(r)/\rho(0)$, which is independent of occupancy.¹³ The logarithm of $[\rho(r)/\rho(0)]$ is plotted as a function of the distance squares ($\log rr$ plot), and this plot intercepts at the origin of the coordinate system.¹³ Its slope defines the variance of the total ED distribution.¹³ Differences in the slope between the query and reference were least-squares minimized. Once two curves had the same slope, relative ENs between the query and reference were analytically fitted by minimizing the differences in the two curves.¹³ Mn2 of each OEC was initially chosen as an internal reference for this analysis.

RESULTS AND DISCUSSION

Relative Electron Numbers of Fe Centers in Nitrogenase Structures. We applied our analytic method to three nitrogenase crystal structures at a resolution of 1.60 Å (Figure 1), corresponding to a chemically reduced ferrous state (1qgu), a chemically oxidized ferric state (1qh1), and a native (ground) state with mixed ferrous–ferric oxidation states (1qh8).¹⁷ Oxidation state differences are relatively small on the four Fe centers (Fe1–Fe4) in the first half of the P cluster (the remaining half has multiple conformations). The standard deviation of differences between two independent P clusters is observed to vary from 0.003 to 0.022 unit in reference to Fe1 of subunit A of the P cluster (CLF) (Figure 1). Among these Fe centers, Fe4 of the reduced structure is a statistical outlier. When this center is excluded, the standard deviation is <0.010 unit. This standard deviation does not change much when the data are further blurred with a ΔB of 10 Å^2 to lower the effective resolution to $\sim 2 \text{ Å}$ (Figure 1). This small variation in relative ENs among four Fe centers likely corresponds to differences in electronic states of individual Fe centers but is not due to different relative occupancies. This number represents an upper estimate of errors of our analytic procedure when we have ignored the contribution of X-ray radiation-induced modifications of oxidation states of Fe centers during data collection.^{37–41} There is no evidence for varying occupancy of the Fe centers in nitrogenase, unlike some other FeS cluster-containing enzymes in which the occupancy of certain Fe centers may change during the catalytic cycle.⁴²

Unlike dark-adapted PSII crystal structures for which a large number of crystal structures were determined in the same putative oxidation state, no X-ray structure of nitrogenase from the same species has been repeated in a given single oxidation state, making it difficult to judge whether small variations in the relative ENs observed fully correspond to different electronic states of individual Fe centers or to some unidentified experimental errors. For this reason, we assume that all Fe centers in the P clusters have the same relative ENs and we

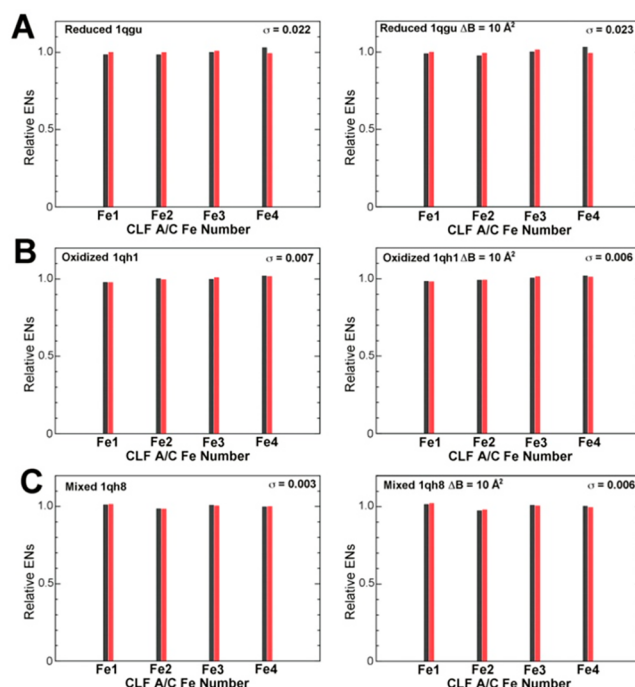


Figure 1. Comparison of relative ENs determined for different Fe centers of a half-P cluster between two independent P clusters (CLF) in three nitrogenase structures at 1.60 Å resolution. (A) Chemically reduced state of the 1qgu structure. (B) Chemically oxidized state of the 1qh1 structure. (C) Native mixed oxidation state of the 1qh8 structure. Left column: results of analysis on the original $F_0 - F_c$ omit maps. Right column: $\Delta B = 10 \text{ Å}^2$ was added in the omit $F_0 - F_c$ maps. Subunits A (black) and C (red) of four CLF Fe centers (1–4) are plotted. See Figure S3 for a close-up of the plot.

calculate standard deviations from the mean value as summarized in Table S1 with a maximum of 0.028. This standard deviation is in general slightly larger than the pairwise difference of the same Fe center between the two independent P clusters (Table S1). It should be noted that delocalized electrons are completely invisible in our analysis because we carry out spherical averaging within a small radius, typically approximately one-half to three-quarters of coordination bond lengths.

Because the resolution of two other nitrogenase structures is at 1.00 and 1.08 Å, we have systematically examined the accuracy of analytically fitted parameters as a function of resolution in a large resolution range.^{15,16} At such high resolution, an approximation of a single Gaussian function for the ED peaks begins to break down, and a $\log rr$ plot is no longer a straight line. However, the analytic method still appears to work reasonably well as theory predicts. We further apply B-factor blurring effects to reduce the effective resolution of these two structures to $\sim 2.0 \text{ Å}$, corresponding to the resolution of the PSII data sets. This analysis and several others consistently show high accuracy of relative ENs to $\sim 1\%$ at $\sim 2.0 \text{ Å}$ resolution (see Figures S1–S4).

Estimation of Uncertainty of Relative Electron Numbers in PSII Crystal Structures. Using five nitrogenase crystal structures, we establish the range of experimental errors on relative ENs determined using our analytic fitting procedure where we assume that all Fe metal ions in the P cluster have the same approximate relative ENs.^{15–17} Next, we estimated the computational uncertainty of our procedure (i.e., precision indexes) using atomic parameters of the 6jly PSII structure.¹⁸ Precision indices are lower bound estimates of errors and upper

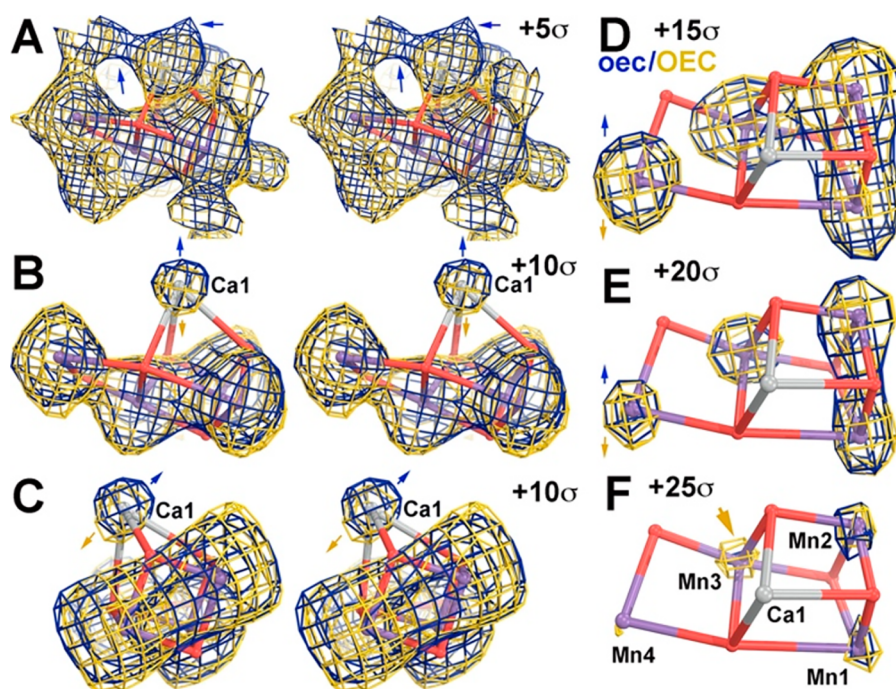


Figure 2. Alignment of the σ_A -weighted $2F_o - F_c$ map for the two OECs of the 6jlj structure contoured at $+5\sigma$ (A) and $+10\sigma$ (B and C in two views), $+15\sigma$ (D), $+20\sigma$ (E), and $+25\sigma$ (F). Arrows indicate the directions of atomic displacements for Ca1 and Mn4 (D). Panels A–C are stereodiamonds.

bound estimates of the accuracy of measured diffraction data. They represent the actual measurement errors in the absence of systematic errors [which are partly canceled out (see the discussion in the Supporting Information)]. Likewise, the reproducibility of the relative ENs determined (using fitting of the two parameters of the EN and the variance of the ED distribution relative to an internal Mn2 reference as described below) represents the uncertainty or error of the values obtained in the absence of systematic errors.

To estimate the computational uncertainty of relative ENs determined using our analytic procedure, the same procedure was applied to a synthetic OEC data set, which was generated at 2.0 Å resolution with the 6jlj coordinates after all translation–libration–screw (TLS) parameters were removed.¹⁸ Standard X-ray form factors of neutral atoms (with four Gaussian terms and one constant^{25–28}) were used so that all Mn centers had 25 electrons of neutral Mn atoms. During Fourier inversion, both OECs and their ligands were fixed in place and their atomic *B*-factors were replaced by their mean value of 43.30 Å², whereas the remaining coordinates were shifted randomly with a $|\Delta r|$ of 0.178 Å using the program CNS.⁴³ This introduced non-systematic random errors of 21.3% into the simulated “observed F_o ” amplitudes of the synthetic data set of the 6jlj coordinates at 2.0 Å resolution.¹⁸ Using this data set, peak maxima in the resulting σ_A -weighted $2F_o - F_c$ map vary slightly among the eight Mn centers due to a different extent of partially overlapping ED distributions of their O ligands. In the σ_A -weighted $F_o - F_c$ maps, all eight Mn centers have the same peak maximum and the same distribution of ED peaks, suggesting that the full peak decomposition has been achieved using these omit maps.

Our analysis shows that the relative ENs determined using this synthetic OEC data set were fully reproduced for each of the four Mn centers in the OEC. In this simulation, all four Mn centers suffered from the same extent of effects of Fourier series termination so that the errors were completely canceled out. In a second simulation test, the *B*-factor for Mn1 of the second OEC

was reduced by a ΔB of -5 Å^2 and that of Mn4 was increased by a ΔB of 5 Å^2 , so that these two Mn centers had a different missing content of information beyond the resolution cutoff of 2.0 Å from that of an internal reference Mn2 center. Our analytic fitting could reproduce these ΔB values within all three significant numbers and the total relative ENs to the fourth significant figure (~ 0.001 unit or 0.02 electron). Therefore, with a small variation range of atomic *B*-factors, the subtle different extent of ripple effects of Fourier series termination is largely invisible.

In addition to the two simulations for the 6jlj structure mentioned above,¹⁸ four types of possible experimental errors that might affect the accuracy of relative ENs determined are discussed further in the Supporting Information and Figures S2 and S3. We want to emphasize that the estimated errors using six specific analyses for PSII and one inference from the FeMo nitrogenase structures may also have a large error in them because we do not know the precise results for accurate determination of errors and because other types of systematic errors could still exist in diffraction data. Therefore, the estimated error may represent the precision index of the methods described in this study, which may represent a lower bound of the actual error.

Different Geometry between Two OECs of a PSII Dimer. Figure 2 shows the alignment of σ_A -weighted $2F_o - F_c$ ED maps for the two OECs in 6jlj, superimposed with the coordinates of the oec (OEC of the lowercase monomer as defined in the given coordinate file).¹⁸ Blue and gold surfaces in Figure 2 represent the ED of the OECs for the lowercase monomer (oec) and uppercase monomer (OEC), respectively. The positions of Ca1 and Mn4 in these densities are not the same between the two OECs. The density of Mn3 of the OEC is higher than that of Mn3 of the oec (Figure 2F).

The analysis of the other four dark-adapted structures of PSII dimers also reveals that the two OECs of each of those structures have different geometries, which is fully consistent with the

conclusions of Kamiya and colleagues.¹⁰ These observations suggest that a significant fraction of the two OECs of dark-adapted PSII structures have oxidation states different from the presumed S_1 state, or different occupancies of metal ions. For example, we show below that there is a large unexplained residual density peak next to the Ca^{2+} ion of the oec in many dark-adapted structures of PSII, highlighting the possibility of different OEC structures. An analysis of this large unexplained peak is beyond the scope of this study and will be described elsewhere. The differences observed here question the validity of the common assumption that the OECs in dark-adapted PSII crystal structures have the same metal ion composition and are fully synchronized in the S_1 state.

Quantitative Analysis of Omit Electron Density Peaks of an OEC. Figure 3 shows the metal ion-omit ED maps used for

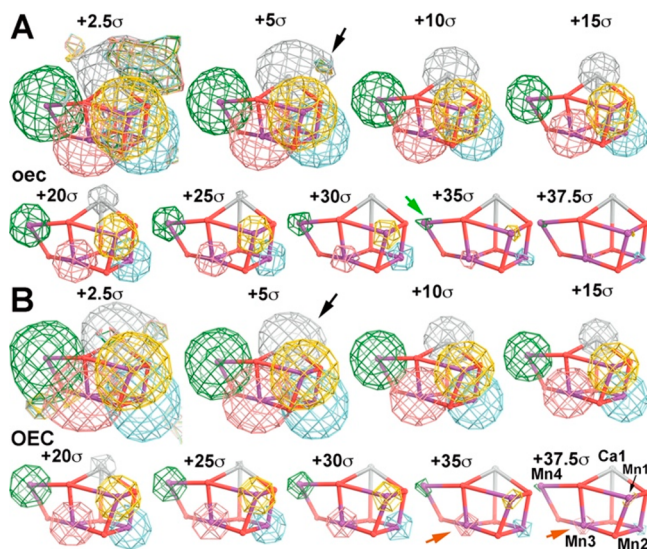


Figure 3. Metal ion-omit maps for the two OECs of the 6jll structure, contoured at $+2.5\sigma$, $+5\sigma$, $+10\sigma$, $+15\sigma$, $+20\sigma$, $+25\sigma$, $+30\sigma$, $+35\sigma$, and $+37.5\sigma$. (A) Lowercase subunit oec. (B) Uppercase subunit OEC. The green arrow indicates the Mn4 peak with the lowest ED peak of the oec. The orange arrows indicate the Mn3 peak with the highest ED value of the OEC. The black arrows highlight difference features between the two OECs next to Ca1. Individual metal ions are colored differently in the omit $F_o - F_c$ maps.

quantification of the ENs of the Mn centers of the two OECs of the 6jll structure.¹⁸ In the OEC, the density of Mn3 is higher than those of the other three Mn centers (orange arrows in Figures 2F and 3B) while Mn4 has the lowest density. In the oec, however, the density of Mn2 is higher than those of the other three Mn centers (Figure 2A). The ED also shows that the Mn peaks of the omit maps are largely spherically symmetric for contour levels above $+5\sigma$ (Figure 3). There are only minor differences in the ED plots with the exception of Mn1 in the oec that is slightly elongated, asymmetrically toward O3. However, we note that there is an unexplained ED peak next to Ca1 in the atomic model of the oec, which is clearly visible at the $+5\sigma$ contour level (black arrow in Figure 3A). A similar peak is much smaller in the OEC (black arrow in Figure 3B), being barely visible only at $+2.5\sigma$. The asymmetric nature of this peak is observed in nearly all high-resolution crystal structures of dark-adapted PSII dimers with some variation, including the four other XFEL structures analyzed here.^{11,18–20}

After fitting the variances of the ED distribution analytically between query and reference atoms, we determined the relative ENs (Figure 4). We selected Mn2 as a reference for each OEC

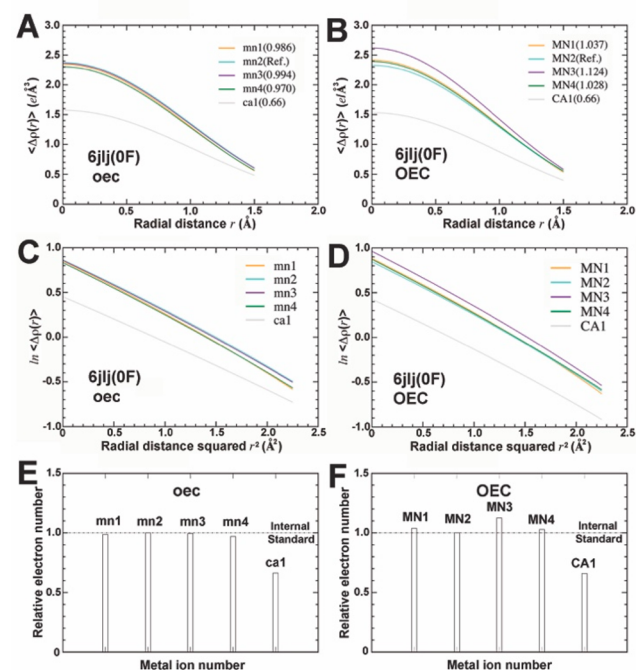


Figure 4. Density plots for the two OECs of the 6jll structure. (A and B) Spherically averaged ED plots as functions of the radial distance from each metal center of the oec and OEC, respectively. (C and D) Logarithms of ED as a function of the radial distance squared (\AA^2). (E and F) Bar graph representations of ENs relative to Mn2 for each Mn and Ca ion. See Figures S5–S8 for the other four PSII structures.

because it exhibits the most spherically symmetric ED peak and has the smallest residual $F_o - F_c$ features nearby in all of these structures. After fitting, the relative ENs were rescaled to the mean EN value of the four Mn ions ($\langle \text{Mn} \rangle$ or $\langle \text{EN} \rangle$) of each OEC. Because the rescaling was applied to each of the metal ions within an OEC, it did not alter the relative ratios among them. Table 1 shows the percentage differences of individual Mn ions relative to $\langle \text{Mn} \rangle$ of each OEC. Table 1 also summarizes the

Table 1. Fractional Variation of ENs of Mn Ions Relative to the Average Value $\langle \text{Mn} \rangle$ within the OEC

$\langle \text{Mn} \rangle$ electron difference (%) between the oec and OEC					
6jll	6dhe	6w1o	4ub6	4ub8	
2.0	2.8	8.4	0.1	0.1	
variation from $\langle \text{Mn} \rangle$ (%)					
	Mn1	Mn2	Mn3	Mn4	Ca1/ $\langle \text{Mn} \rangle$
6jll-oec	−0.1	1.3	0.7	−1.8	67.2
6jll-OEC	−1.0	−4.5	7.3	−1.9	62.9
6dhe-oec	3.5	7.3	−2.2	−8.6	75.0
6dhe-OEC	3.1	1.8	4.0	−8.8	70.1
6w1o-oec	1.2	8.0	−0.6	−8.6	72.6
6w1o-OEC	4.7	0.6	1.9	−7.2	68.7
4ub6-oec	−0.6	0.5	1.4	−1.3	53.8
4ub6-OEC	3.8	−2.4	2.7	−4.1	60.9
4ub8-oec	−5.4	2.9	0.8	1.7	50.4
4ub8-OEC	2.5	−2.0	2.9	−3.5	56.8

corresponding analysis for the 6dhe, 6w1o, 4ub6, and 4ub8 XFEL structures (Figures S5–S8).^{11,18–20} Like the 6jlj structure, the four other structures have different ED distributions for the Mn ions relative to one another in the two OECs of each dimer, suggesting differences in either occupancy, redox state, or both.^{11,18–20}

Variation of Relative Electron Numbers in Metal Ions of the OEC. We determine the relative ENs of the Mn ions in each OEC to explore the origin of the variations between the two OECs in each dark-adapted PSII dimer (Table 1). The two OECs could be in two different oxidation states if each OEC corresponds to a single redox state or in different average oxidation states if each OEC represents a mixture of different oxidation states. Most of the OEC pairs exhibit significant differences, with the largest difference of 8.4% observed in the 6w1o structure (Table 1).^{11,18–20} Also, although the two OECs of the 4ub6 or 4ub8 structure have a similar mean value of the EN, the distribution of relative ENs among the four Mn ions in each OEC is different. Specifically, the percentage deviations relative to $\langle \text{Mn} \rangle$ in the oec of 4ub6 are -0.6 , 0.5 , 1.4 , and -1.3 for Mn1–Mn4, respectively, and 3.8 , -2.4 , 2.7 , and -4.1 for the corresponding metal centers in the OEC (Table 1).⁷ In summary, none of the five crystal structures have two OECs with equivalent distributions of ENs for the four Mn ions (Table 1).^{11,18–20}

The differences observed here are much larger than the experimental uncertainty estimated using nitrogenase structures described above. Differences in the relative ENs of Mn ions when comparing the two OECs of a dimer can be interpreted as resulting from OECs in different oxidation states, particularly when the two OECs have the same occupancies of metal ions. Equal occupancies, however, are usually assumed but not validated. To deconvolute the contributions from varying occupancies and varying oxidation states, we consider the extent of variation of relative ENs among the four Mn ions in the S_1 state.⁹ The oxidation states of the four Mn ions in the S_1 state are most commonly assigned as $+3$, $+4$, $+4$, $+3$ or (3443). Therefore, the mean oxidation state ($\langle \text{oxidation} \rangle$ or $\langle \text{Mn} \rangle$ in either the EN unit or the oxidation unit) is 3.5 , and the $\langle \text{EN} \rangle$ value is 21.5 electrons.⁹ The EN percentage difference between Mn^{4+} (21 electrons) and Mn^{3+} (22 electrons) is $\sim 4.6\%$, relative to each other, and the expected percentage deviations relative to the $\langle \text{EN} \rangle$ are 2.3% , -2.3% , -2.3% , and 2.3% for the four Mn ions. The comparison to data in Table 1 shows that none of the OECs from the five analyzed structures exhibits such a distribution expected for the S_1 state. Many of the observed differences far exceed the variation of values expected for the S_1 state, suggesting that the occupancies of some metal ions [i.e., Mn4 and Ca^{2+} ions (see the discussion below)] are likely to vary and to be lower than those of the remaining metal centers of the OEC. In addition, the S_1 model with Mn2(IV) and Mn3(IV) is expected to have the same ENs of Mn2 and Mn3.²⁰ However, our analysis shows that none of the structures has such a distribution (Table 1),^{11,18–20} suggesting that the redox state of the OEC cannot be uniformly assigned to the fully synchronized S_1 (3443) state as expected for the dark-adapted state of PSII.

Reduced Ca^{2+} Occupancy in the OEC. The theoretical EN ratio between the Ca^{2+} ion (18 electrons) and $\langle \text{EN} \rangle$ (21.5 electrons) of the OEC in the S_1 state is 83.7% .⁹ In our analysis, however, the highest ratio observed is 75.0% in the oec of 6dhe and the lowest ratio is 50.4% in the oec of 4ub6 (Table 1).^{11,19} Even if the mean oxidation state is assumed to be $+3.0$ (instead of $+3.5$), which conflicts with the $+3$, $+4$, $+4$, $+3$ oxidation state

of S_1 ,⁹ the theoretical ratio is 81.8% , which is still well above the observed ratios. These observations suggest that the occupancy of Ca^{2+} in these OECs varies from 60% to 90% , assuming the mean oxidation state of the OEC is $+3.5$. Notably, the dark-adapted structures of *Thermosynechococcus vulcanus* PSII dimers (6jlj, 4ub6, and 4ub8) have a Ca^{2+} ratio relative to $\langle \text{EN} \rangle$ lower than that of the *Thermosynechococcus elongatus* PSII dimer by as much as $\sim 10\%$ (6dhe and 6w1o) (Table 1).^{19,20}

Overall, the EN variation of Ca^{2+} is between 7% and 12% when comparing the two OECs of each PSII dimer in the five structures (Table 1),^{11,18–20} significantly higher than the expected value associated with the different mean oxidation states of the two OECs and much larger than the maximal uncertainty estimated for the relative ENs determined. Thus, the two OECs of all *T. elongatus* and *T. vulcanus* PSII dimers in these crystal structures likely have a substoichiometric Ca^{2+} occupancy. Below we provide independent crystallographic evidence of an abnormal distribution of refined atomic B -factors of Ca^{2+} ions of the OEC that indicates the presence of a fraction of Ca-depleted OEC in crystals. The Ca1-depleted OEC can be ignored in a kinetic oxygen evolution assay because it does not generate oxygen and thus is completely invisible. However, it is highly visible in electron density maps according to the analysis described herein.

Independent Crystallographic Evidence of the Ca-Depleted OEC. Atomic B -factors between chemically bonded and indirectly connected atoms through bonding with the same atom (also known as angularly bonded atoms) in structures do not vary as much and are often restrained geometrically during structural refinement because atomic motions of bonded atoms are highly correlated ($B = 8\pi^2|\Delta r|^2$). When large B -factor differences are observed between bonded atoms within the OEC in the presence of strong B -factor restraints, it is almost certain that the composition of atoms in the OEC is incorrect because the uniform occupancy assumption of the OEC is not valid. For example, the B -factor of Ca1 of the lowercase OEC in the 4ub6 structure is 34.04 \AA^2 , and the mean value for the remaining atoms of the OEC is $21.48 \pm 2.09 \text{ \AA}^2$ (Figure S8A).¹¹ This B -factor value of Ca1 is 6 times the standard deviation above the mean value and a significant outlier. It is shown in the Supporting Information that the variance values of atomic motions for all metal ions of the OEC in metal ion-omit ED peaks in that structure are almost identical, and that Ca1 has a much-reduced occupancy relative to that of the Mn ions (Figure S8). Similarly, Ca1 of the oec of the 4ub8 structure is also a significant outlier (Figure S8B).¹¹

Different distributions of refined B -factors for other atoms of the OECs between the two OECs of a given structure have also revealed that the two OECs may be in different oxidation states, i.e., between the two structures of 4ub6 and 4ub8.¹¹ The refined B -factor of the O3 atom of the oec of the 4ub8 structure is a significant outlier (its $B = 35.77 \text{ \AA}^2$, and the mean value of that OEC $\langle B \rangle = 23.79 \pm 2.35 \text{ \AA}^2$), with its value being 5.1 times the standard deviation above the mean value. This is likely due to the reduced occupancy of this O ligand. This outlier does not occur in the second OEC of the same structure. If the number of O ligands inside the OEC is dependent on the mean oxidation state of the OEC, the different occupancy of O3 of the OEC in the two OECs likely indicates that they are in a different oxidation state. Determining relative occupancies of O ligands of the OEC as a function of flash number in light-dependent reactions is critical in order to establish the $\text{O}=\text{O}$ bond formation

mechanisms catalyzed by this enzyme. This issue will be further discussed in a future publication.

The structural biology community has often misused *B*-factors as fudge factors for model refinement under invalid assumptions. Whenever invalid assumptions were made or invalid models were built for model refinement, the resulting ED maps could have significant assumption-based or model-based biases, which could sometimes lead to overinterpretations of structural data.^{13,14,38–40,44–57} For some time, we were very concerned about the validity of some earlier data processing procedures of XFEL data sets, the quality of processed data, structures being derived, and overinterpretation of some XFEL data for PSII intermediates.^{38–40,54,56,57} It was often assumed that the occupancy and *B*-factor could not be simultaneously refined in crystal structures determined at ~ 2.0 Å resolution. Because the *B*-factor affects the slope of the ED distribution in the log *rr* plot and occupancy affects the entirety of the ED distribution curves linearly, the two parameters can be accurately decoupled in omit maps using the methods described in this study. Thus, this study provides new methods for extracting new information from diffraction data of macromolecular crystal structures.

Reduced Mn4 Occupancy in the OEC of PSII from *T. elongatus*. The OECs of two dark-adapted structures from *T. elongatus* PSII dimers (6dhe and 6w1o) exhibit ENs of Mn4 that are 7.2% and 8.8% lower than the $\langle \text{Mn} \rangle$ value (Table 1).^{19,20} These values are 3.1 and 3.8 times lower than the difference corresponding to one oxidation unit (2.3%).⁹ Therefore, these large differences could likely be due to a lower occupancy of the Mn4 site in the OEC from *T. elongatus*. In contrast, OECs from *T. vulcanus* PSII dimers (6jlj, 4ub6, and 4ub8) do not exhibit this feature, with ENs of the four Mn ions relative to $\langle \text{Mn} \rangle$ within the expectation of one or fewer oxidation units (Table 1).^{11,18–20} Whether a reduced occupancy of Mn4 of the OEC in these structures should be correlated with the reduced electron numbers of its oxygen ligands will be addressed elsewhere. It is conceivable that oxidative decarboxylation of Mn4 ligands (induced by X-ray radiation) is responsible for its reduced occupancy.⁵⁸ Occupancy variation of Mn ions of the OEC can also be corroborated by unusual distributions of refined atomic *B*-factors in many structures, which will be systematically described and discussed elsewhere.

Even after Mn4 is excluded from the calculation of $\langle \text{Mn} \rangle$, the relative ENs of Mn obtained in the two OECs of PSII dimers from *T. elongatus* suggest that the OECs are in different oxidation states.^{19,20} For example, the OEC of 6w1o has relative ENs of 4.7%, 0.9%, and 1.9% for Mn1–Mn3, respectively, in percentage deviations relative to $\langle \text{Mn} \rangle$ (Table 1).²⁰ That distribution is close to the expected oxidation states of +3, +4, and +4 for these three centers in the S_1 state.⁹ However, the oec of the same structure has a completely different distribution of 1.2%, 8.0%, and -0.6% in percentage deviations relative to $\langle \text{Mn} \rangle$,²⁰ suggesting that the oec might be in a different oxidation state, assuming that their relative occupancies for these three high-affinity Mn ions are the same. At the resolution of these structures, deconvolution of the relative occupancy and the number of valence electrons remains very difficult. It requires the further development of methods in both real space and reciprocal space for quantitative analysis of relative electron numbers and occupancy of Mn ligand atoms.

Minimization of Atomic Scattering Factor Model Bias in Omit Electron Density Maps. Model refinement is a method for deconvoluting ED peaks into individual atoms that

relies on prior knowledge of atomic scattering factors and known oxidation states of metal ions. The substitution of X-ray atomic scattering factors of neutral atoms for ionic species for calculation of both structure factors and ED maps in the crystallography of macromolecules will result in an overlooked model bias of input atomic scattering factors. Most crystallographic programs automatically carry out this substitution, which will introduce an error of four electrons for a Mn^{4+} ion (i.e., as much as 20% in $25-21 e$ versus $21 e$ in its relative EN; furthermore, no information about atomic scattering factors is available for a Mn^{4+} ion²⁸). The error introduced by this substitution is proportional to the magnitudes of atomic scattering factors. Thus, it affects the ENs of ligand atoms more than that of the Mn ion center. Ligands have fewer electrons than metal ions, and the tail functions of ligands at the Mn centers are smaller than the tail function of metal ions at the ligand centers. Additionally, the resulting atomic *B*-factors obtained using such model refinement are significantly affected by this substitution and no longer represent the variance of atomic motions. Thus, the intrinsic variance of atomic ED functions and the variance of atomic motions are completely merged together.

In this study, we do not apply any prior knowledge of oxidation states of metal ions and their unknown atomic scattering factors in our analysis. We omit metal ions of interest with unknown oxidation states. Our analysis includes two parameters, the total EN inside the ED peak and the variance of the ED distribution within a given radius of the peak. At a resolution of ~ 2.0 Å, the variance of the ED distribution is not separable, and one need not separate atomic motions and the intrinsic variance of the atomic ED peak of a stationary atom for estimation of relative ENs. Moreover, when the Mn ion is spherically averaged in polar coordinates, the error is propagated through a Fourier–Bessel transformation and becomes even smaller.⁵⁹ The resulting error is thus only a minor deviation of the ED distribution of the Mn ions from a spherically symmetric distribution, which becomes largely invisible in the spherically averaged ED plots for the structures analyzed here.

Applicability and Limitations of Our Analytic Method.

We want to emphasize that the analytic method described in this study can be applied only to crystal structures of large macromolecules to determine the relative ENs of metal ions, but not to those of small molecules. This method is not suitable for determination of the absolute ENs in reference to the calculated structure factors because it does not carry out an integration of the entire EN peaks over the entire space. Instead, it fits analytically within an ~ 1.0 Å radius of a sphere where semilogarithms of plots of the spherically averaged density as a function of distance squared exhibit a linear relationship. This method can also become unreliable when the *B*-factor difference between the reference and query metal ions is very large because different extents of Fourier series truncation at the highest-resolution shell are not corrected in the current procedure. This is no longer a problem when both the reference and queries are within the same OEC. The relative ENs determined cannot easily be decomposed into the change in relative occupancy and the change in oxidation state. However, we can provide some estimates of contributions from these two factors because the change in relative oxidation state has a much smaller range of relative ENs than the change in relative occupancy.

The valence electrons are present in outer shells of atoms where transfer of valence electrons between bonded atoms occurs. Thus, atomic scattering factors differ primarily at

medium and low resolution between neutral atoms and ionized atoms. For example, there are 20% differences in ENs between a neutral Mn atom and a Mn^{4+} ion, the largest differences in their atomic scattering factors between the two Mn species, but their atomic scattering factor differences are primarily located below 2.0 Å resolution, including 20% at zero scattering angle $s = 0$. With an increase in the level of atomic motion and a decrease in the measurable resolution of diffraction data, differences between them are larger and can be measured accurately beyond 2.0 Å resolution. Thus, information about valence electrons and their transfer is highly enriched in diffraction data of very large macromolecules due to the high density of oversampling in reciprocal space at relatively medium and low resolution. In the crystal structure of PSII at ~ 2.0 Å resolution, there are approximately 500000 Bragg reflections and all of them encode the enriched information about the ENs of Mn ions.⁸ For this reason, substitution of neutral atoms for ionized metal ions and protein atoms can introduce large modeling errors into the resulting electron density. In this study, omit electron density peaks were used for analysis and thus modeling errors of atomic scattering factors have been minimized. For comparison, the same substitution of atomic scattering factors has no apparent effect on the crystal structure of MnO. In fact, it affects only a single Bragg reflection within the 2.0 Å resolution limit in that structure.²²

Our method differs from the classic integration of the radial distribution function for accurate determination of the absolute ENs for the metal ion in real space.²² We cannot apply *B*-factor blurring to these structures to match those of very large macromolecules because if we do so, we have only a single Bragg reflection below 2.0 Å that still contains the information about the electron number. In fact, reciprocal space multipolar model refinement in general could not result in a reliable number of valence electrons, either, for crystal structures even at 0.37 Å resolution.²⁹ An analysis of the deviation of multipolar model refined structure factors from spherically averaged amplitudes for carbonyl O atoms at 0.37 Å resolution²⁹ using the method described in this study shows that the largest aspherical deviations are near 3.0 Å resolution and that the deviation largely disappears at 2.0 Å resolution. This is distinctively different from anisotropic motions of atoms.

Physical Meaning of Crystallographic Electron Numbers (ENs). What do relative ENs determined in this study mean physically? Because our method is reliant on spherical averaging and limited within a radius of ~ 1.0 Å, the resulting relative ENs are closely related to single-pole point charges used in molecular dynamics simulations, but not multipole charges. Moreover, they do not contain information about the delocalized valence electrons or dispersion and anomalous electrons.

Crystallographic charge numbers determined from ED maps in general are differences between the absolute ENs integrated from ED peaks and their atomic number *Z*. They may differ from formal charges. The absolute ENs integrated radially from ED peaks for Mg(II), Mn(II), Fe(II), Co(II), and Ni(II) in crystal structures of MgO, MnO, FeO, CoO, and NiO show that the actual crystallographic charge numbers are +1.9, +1.4, +1.3, +1.2, and +0.7 electrons, respectively, which differ by 0.1–1.3 electrons from charges derived from formal oxidation states.^{22,23} These differences can be explained using different extents of covalency in these ionic structures.^{22,23} For the crystal structures of small molecules (i.e., the CaF_2 crystal⁶⁰) at the resolution corresponding to the wavelength of the Ag radiation source, the

topology and the boundary of atoms are well-defined for which the absolute ENs can be determined using real-space integration, and the absolute number determined this way still differs from radial integration after spherical averaging when the diffraction data are filtered to the resolution of the wavelength of Cu and Mo radiation sources.

How are the relative ENs determined in this study related to those with other methods accessible with comparable resolution? Our analysis does not include corrections for dispersion and anomalous electrons because diffraction data were collected away from K-edge energies of the metal ions analyzed. In fact, crystallographers began to develop spatially resolved dispersion/anomalous diffraction methods to allow assignment of oxidation states of metal ions.⁴³ This assignment is based on the assumption that metal ions inside the multicentered metal ion cluster behave the same as or similar to those of single-center metal ions. However, the validity of this assumption has recently been challenged in various synthetic multicenter Fe cluster compounds in which both spectroscopic data and dispersion diffraction data were measured.⁶¹

Quality of Atomic Models and Its Effect on the Accuracy of Relative Electron Numbers Determined.

The quality of the atomic models used for calculation of omit ED maps does affect the accuracy of the absolute values of ENs determined, more so for the absolute ENs than for the relative ENs. For small molecules, model *R*-factors of crystal structures are refined to 1–2% at subangstrom resolution.^{22,23,29,60} However, model *R*-factors for protein structures are much larger. For example, the cross-validation free *R*-factor for the 6jll structure is 19.8% at 2.15 Å resolution, while for the 6wlo structure, it is 24.0% at 2.08 Å resolution.^{18,20} Therefore, the results obtained for the 6jll structure should be more accurate and more reliable than those for 6wlo.^{18,20} Here, we use one of the metal ions (Mn2 of each OEC) as an internal reference, which improves greatly the accuracy and reliability of relative ENs because many modeling errors tend to cancel out (see the [Supporting Information](#)). For example, missing ordered water molecules or lipids in the atomic models distant from the OEC can alter the ED distributions in the omit ED peaks for both query and reference Mn ions, and these errors would prevent determination of their absolute ENs. However, their effects are largely canceled in our relative fitting because they affect similarly the unknown and reference metal ions (see the [discussion in the Supporting Information](#)). Similarly, the effects of anisotropy on the ED distributions are also largely canceled out because the tensors of anisotropy are often aligned in the same way within each OEC.

Biochemical Implications of Nonsynchronized OECs.

Our quantitative analysis of the omit ED maps shows that 10–40% of the OECs are Ca^{2+} depleted, which amount varies in XFEL diffraction data from different crystal structures ([Table 1](#)). At femtosecond duration, the nuclear positions obtained should remain unchanged and be thought to reflect those before any radiation damage. However, recent computer simulations do not support this hypothesis for many of the currently used X-ray pulse lengths, making proper interpretation of XFEL diffraction data a challenge.⁶² XFEL radiation-induced changes in the ED distribution may be very significant.^{39,40,63} If their effects can be assumed to be independent of their locations, the reduced occupancies of metal ions should also represent those in the crystal before irradiation and thus likely be associated with sample purification and crystal preparation.

The Mn contents of purified the *T. elongatus* monomer and dimer were estimated to be ~ 3.8 and ~ 3.7 Mn atoms/36 chlorophylls, respectively, the latter of which was $\sim 93\%$ of those of the four expected Mn ions,⁶⁴ assuming that there was no Mn oxide nanoparticles in the solution. This implies that if one OEC has 100% occupancy, the second one could have as low as 86% occupancy with a 14% difference between the two OECs of each PSII dimer. This is clearly within the experimentally detectable limit of the methods described here. Interestingly, we observed here that *T. elongatus* PSII XFEL structures have a content of the apparent Mn4-depleted OEC higher than that of *T. vulcanus* PSII (Table 1).^{11,18–20} It appears that the Mn4-depleted OECs represent ~ 7 – 12% of the centers with an upper bound estimation of the differences in overall occupancies between the two monomers in crystals of $\leq 8.4\%$ (Table 1), which is fully consistent with the estimated occupancy of the complete OEC in purified PSII samples.⁶⁴ The variation of the Mn occupancies in these XFEL structures analyzed here is also supported by usually large deviations of atomic *B*-factors refined to the mean values of the OEC, as discussed above for the Ca^{2+} ion.

Moreover, as much as one-third of the crystalline PSII samples had to be discarded because of the very high content of Mn^{2+} ions.⁶⁵ Why some preparations of PSII crystalline samples contained a low content of Mn^{2+} ion and why others had a high content remains unexplained. This might be related to some instability of the OEC in purified PSII samples, which could lose a significant amount of the oxygen evolution activity during the long time scale of crystallization.⁶⁶

Loss of Ca^{2+} from the OEC is well-known in purified plant PSII membranes and has been extensively studied, often associated with loss of two Ca^{2+} -stabilizing extrinsic subunits (PsbP and PsbQ).^{67,68} Heterogeneity of OECs in PSII found *in vivo* is also well-known because PSII is continuously disassembled and reassembled, with the half-life of a fully assembled PSII complex being ~ 20 min under standard light conditions.^{69–71} The extent of fully assembled OECs present in a PSII crystal is generally considered to be higher than in the native environment, although this needs to be quantitatively assessed. Therefore, our study represents the first quantitative assessment of the stoichiometric amount of Ca^{2+} in the OEC in PSII crystals.

Ca^{2+} and Mn4 of the OEC are the only two metal ions that have two terminal water molecule/hydroxide ligands.⁸ Because terminal water/hydroxide ligands could be rapidly exchanged with bulk water molecules, they might not directly contribute to the binding affinity of their corresponding metal centers for the enzyme significantly. This would explain why these metal ions are partially lost in the OECs of PSII dimers in crystals. Alternatively, oxidative decarboxylation of some Mn4 ligands may occur,⁵⁸ which could lower the binding affinity of the Mn4 ion. This requires additional analysis.

Our crystallographic analysis of PSII dimers shows that incomplete OECs are not evenly distributed between the two monomers of PSII dimers and that the mean oxidation numbers of the two OECs are also different (Table 1).^{11,18–20} It is possible that the crystalline environment plays a role in the asymmetry of the two monomers of PSII dimers. However, whether this asymmetry is an intrinsic property of PSII dimers *in vivo* remains to be established. Even if the OECs of dark-adapted PSII were fully synchronized in S_1 before crystallization, many slow light-independent reactions could occur during the long incubation for crystallization, which may also result in asynchronous *S* states in crystals. Thus, the kinetics of a single

photochemical event in a crystal may contain larger stochastic fluctuations than in solution.⁷² Moreover, PSII dimers *in vivo* may contain Ca^{2+} -depleted OECs, which undergo different transitions from the Kok cycle during photochemical reactions given the well-known fact that the Ca^{2+} ion is an integral part of the functional OEC in water oxidation.⁶⁸

In summary, we have shown that crystallographically derived electron density maps contain quantitative information about the electron numbers and variance of the electron density distribution per metal ion peak when their electron density peaks are resolvable. We found that the OECs of each monomer in the PSII dimer are not fully synchronized in the crystallographic dark-adapted “ S_1 state”. Therefore, it is natural to expect that the redox states of the two OECs would not be synchronized during light phase redox reaction, causing structural disorder that would complicate the proper interpretation of spectroscopic data.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biochem.1c00611>.

Three supporting sections, one supporting table, three supporting references, and eight supporting figures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jimin Wang – Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114, United States; orcid.org/0000-0002-4504-8038; Email: jimin.wang@yale.edu

Authors

Christopher J. Gisriel – Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States

Krystle Reiss – Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States

Hao-Li Huang – Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States

William H. Armstrong – Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467, United States

Gary W. Brudvig – Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114, United States; Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States; orcid.org/0000-0002-7040-1892

Victor S. Batista – Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States; orcid.org/0000-0002-3262-1237

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.biochem.1c00611>

Author Contributions

J.W. designed and executed the experiments and wrote the draft of the manuscript. All authors were involved in the analysis and interpretation of results and writing of the manuscript.

Funding

This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of

Chemical Sciences, Geosciences, and Biosciences, via Grants DESC0001423 (V.S.B.) and DE-FG0205ER15646 (G.W.B.).

Notes

The authors declare no competing financial interest.

^{||}Retired.

ACKNOWLEDGMENTS

The authors thank Professors Peter B. Moore and Richard J. Debus for insightful discussion during course of this study.

ABBREVIATIONS

PSII, photosystem II; OEC, oxygen-evolving complex; oec, OEC in lowercase monomer A; OEC, OEC in uppercase monomer B; ED, electron density; EN, electron number per atom.

REFERENCES

- (1) Joliot, P.; Kok, B. Oxygen evolution in photosynthesis. In *Bioenergetics of Photosynthesis*; Academic Press: New York, 1975; pp 388–413.
- (2) Joliot, P.; Barbieri, G.; Chabaud, R. A new model of photochemical centers in system-2. *Photochem. Photobiol.* **1969**, *10*, 309–329.
- (3) Joliot, P. Kinetic studies of photosystem 2 in photosynthesis. *Photochem. Photobiol.* **1968**, *8*, 451–463.
- (4) Kok, B.; Forbush, B.; McGloin, M. Cooperation of charges in photosynthetic O₂ evolution. 1. A linear 4-step mechanism. *Photochem. Photobiol.* **1970**, *11*, 457–475.
- (5) McEvoy, J. P.; Brudvig, G. W. Water-splitting chemistry of photosystem II. *Chem. Rev.* **2006**, *106*, 4455–4483.
- (6) Vinyard, D. J.; Brudvig, G. W. Progress toward a molecular mechanism of water oxidation in photosystem II. *Annu. Rev. Phys. Chem.* **2017**, *68*, 101–116.
- (7) Cox, N.; Pantazis, D. A.; Lubitz, W. Current understanding of the mechanism of water oxidation in photosystem II and its relation to XFEL Data. *Annu. Rev. Biochem.* **2020**, *89*, 795–820.
- (8) Umena, Y.; Kawakami, K.; Shen, J. R.; Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **2011**, *473*, 55–60.
- (9) Luber, S.; Rivalta, I.; Umena, Y.; Kawakami, K.; Shen, J. R.; Kamiya, N.; Brudvig, G. W.; Batista, V. S. S₁-state model of the O₂-evolving complex of photosystem II. *Biochemistry* **2011**, *50*, 6308–6311.
- (10) Tanaka, A.; Fukushima, Y.; Kamiya, N. Two different structures of the oxygen-evolving complex in the same polypeptide frameworks of photosystem II. *J. Am. Chem. Soc.* **2017**, *139*, 1718–1721.
- (11) Suga, M.; Akita, F.; Hirata, K.; Ueno, G.; Murakami, H.; Nakajima, Y.; Shimizu, T.; Yamashita, K.; Yamamoto, M.; Ago, H.; Shen, J. R. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* **2015**, *517*, 99–103.
- (12) Wang, J. Determination of chemical identity and occupancy from experimental density maps. *Protein Sci.* **2018**, *27*, 411–420.
- (13) Wang, J.; Smithline, Z. B. Crystallographic evidence for two-metal-ion catalysis in human pol eta. *Protein Sci.* **2019**, *28*, 439–447.
- (14) Wang, J. Crystallographic identification of spontaneous oxidation intermediates and products of protein sulfhydryl groups. *Protein Sci.* **2019**, *28*, 472–477.
- (15) Zhang, L. M.; Morrison, C. N.; Kaiser, J. T.; Rees, D. C. Nitrogenase MoFe protein from *Clostridium pasteurianum* at 1.08 Å resolution: comparison with the *Azotobacter vinelandii* MoFe protein. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *71*, 274–282.
- (16) Spatzal, T.; Aksoyoglu, M.; Zhang, L.; Andrade, S. L.; Schleicher, E.; Weber, S.; Rees, D. C.; Einsle, O. Evidence for interstitial carbon in nitrogenase FeMo cofactor. *Science* **2011**, *334*, 940.
- (17) Mayer, S. M.; Lawson, D. M.; Gormal, C. A.; Roe, S. M.; Smith, B. E. New insights into structure-function relationships in nitrogenase: A 1.6 Å resolution X-ray crystallographic study of *Klebsiella pneumoniae* MoFe-protein. *J. Mol. Biol.* **1999**, *292*, 871–891.
- (18) 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. An oxyl/oxo mechanism for oxygen-oxygen coupling in PSII revealed by an x-ray free-electron laser. *Science* **2019**, *366*, 334–338.
- (19) 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. Structures of the intermediates of Kok's photosynthetic water oxidation clock. *Nature* **2018**, *563*, 421–425.
- (20) 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. Untangling the sequence of events during the S₂ → S₃ transition in photosystem II and implications for the water oxidation mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 12624–12635.
- (21) Agarwal, R. C. A new least-squares refinement technique based on the fast Fourier transformation algorithm. *Acta Crystallogr., Sect. A: Cryst. Phys., Diff., Theor. Gen. Crystallogr.* **1978**, *34*, 791–809.
- (22) Sasaki, S.; Fujino, K.; Takeuchi, Y. X-ray determination of electron-density distributions in oxides MgO, MnO, CoO, and NiO, and atomic scattering factors of their constituent atoms. *Proc. Jpn. Acad., Ser. B* **1979**, *55*, 43–48.
- (23) Sasaki, S.; Fujino, K.; Takeuchi, Y.; Sadanaga, R. On the estimation of atomic charges by the X-ray method for some oxidizes and silicates. *Acta Crystallogr., Sect. A: Cryst. Phys., Diff., Theor. Gen. Crystallogr.* **1980**, *36*, 904–915.
- (24) Moore, P. B. *Visualizing the invisible: imaging techniques for the structural biologist*; Oxford University Press: Oxford, U.K., 2012.
- (25) Cromer, D. T.; Mann, J. B. X-ray scattering factors computed from numerical Hartree-Fock wave functions. *Acta Crystallogr., Sect. A: Cryst. Phys., Diff., Theor. Gen. Crystallogr.* **1968**, *24*, 321–324.
- (26) Fox, A. G.; O'Keefe, M. A.; Tabborn, M. A. Relativistic Hartree-Fock X-ray and electron atomic scattering factors at high angles. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1989**, *45*, 786–793.
- (27) Waasmaier, D.; Kirfel, A. New analytical scattering-factor functions for free atoms and ions. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1995**, *51*, 416–431.
- (28) Prince, E. *International Tables for Crystallography. Vol. C*; Kluwer Academic Publishers: London, 2014.
- (29) Jelsch, C.; Devi, R. N.; Noll, B. C.; Guillot, B.; Samuel, I.; Aubert, E. Aceclofenac and interactions analysis in the crystal and COX protein active site. *J. Mol. Struct.* **2020**, *1205*, 127600.
- (30) Sim, G. A. The Distribution of phase angles for structures containing heavy atoms. 2. A modification of the normal heavy-atom method for non-centrosymmetrical structures. *Acta Crystallogr.* **1959**, *12*, 813–815.
- (31) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240–255.
- (32) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235–242.

- (33) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- (34) Burley, S. K.; Berman, H. M.; Kleywegt, G. J.; Markley, J. L.; Nakamura, H.; Velankar, S. Protein Data Bank (PDB): The single global macromolecular structure archive. *Methods Mol. Biol.* **2017**, *1607*, 627–641.
- (35) Kleywegt, G. J.; Jones, T. A. Halloween ... masks and bones. In *From First Msp to Final Model*; Hubbard, R., Waller, D., Eds.; Warrington, U.K., 1994.
- (36) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
- (37) Yano, J.; Kern, J.; Irrgang, K. D.; Latimer, M. J.; Bergmann, U.; Glatzel, P.; Pushkar, Y.; Biesiadka, J.; Loll, B.; Sauer, K.; Messinger, J.; Zouni, A.; Yachandra, V. K. X-ray damage to the Mn₄Ca complex in single crystals of photosystem II: a case study for metalloprotein crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 12047–12052.
- (38) Wang, J. X-ray radiation-induced addition of oxygen atoms to protein residues. *Protein Sci.* **2016**, *25*, 1407–1419.
- (39) Wang, J. Oxygen additions in serial femtosecond crystallographic protein structures. *Protein Sci.* **2016**, *25*, 1797–1802.
- (40) Wang, J. Destruction-and-diffraction by X-ray free-electron laser. *Protein Sci.* **2016**, *25*, 1585–1592.
- (41) Meents, A.; Gutmann, S.; Wagner, A.; Schulze-Bries, C. Origin and temperature dependence of radiation damage in biological samples at cryogenic temperatures. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1094–1099.
- (42) Shisler, K. A.; Broderick, J. B. Emerging themes in radical SAM chemistry. *Curr. Opin. Struct. Biol.* **2012**, *22*, 701–710.
- (43) Brunger, A. T. Version 1.2 of the crystallography and NMR system. *Nat. Protoc.* **2007**, *2*, 2728–2733.
- (44) Wang, J. A corrected quaternary arrangement of the peptidase HslV and atpase HslU in a cocrystal structure. *J. Struct. Biol.* **2001**, *134*, 15–24.
- (45) Wang, J.; Rho, S. H.; Park, H. H.; Eom, S. H. Correction of X-ray intensities from an HslV-HslU co-crystal containing lattice-translocation defects. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 932–941.
- (46) Wang, J.; Kamtekar, S.; Berman, A. J.; Steitz, T. A. Correction of X-ray intensities from single crystals containing lattice-translocation defects. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 67–74.
- (47) Wang, J. DNA polymerases: Hoogsteen base-pairing in DNA replication? *Nature* **2005**, *437*, E6–7.
- (48) Wang, J. Inclusion of weak high-resolution X-ray data for improvement of a group II intron structure. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 988–1000.
- (49) Wang, J.; Wing, R. A. Diamonds in the rough: a strong case for the inclusion of weak-intensity X-ray diffraction data. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 1491–1497.
- (50) Wang, J. On the validation of crystallographic symmetry and the quality of structures. *Protein Sci.* **2015**, *24*, 621–632.
- (51) Wang, J. Estimation of the quality of refined protein crystal structures. *Protein Sci.* **2015**, *24*, 661–669.
- (52) Wang, J. Comment on “Crystal structures of translocator protein (TSPO) and mutant mimic of a human polymorphism. *Science* **2015**, *350*, 519.
- (53) Wang, J. Misreading chaperone-substrate complexes from random noise. *Nat. Struct. Mol. Biol.* **2018**, *25*, 989–990.
- (54) Wang, J.; Armstrong, W. H.; Batista, V. S. Do the crystallographic XFEL data support binding of a water molecule to the oxygen-evolution complex of photosystem II exposed to two flashes of light? *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2023982118.
- (55) Wang, J.; Natchiar, S. K.; Moore, P. B.; Klaholz, B. P. Identification of Mg(2+) ions next to nucleotides in cryo-EM maps using electrostatic potential maps. *Acta Crystallogr. D Struct. Biol.* **2021**, *77*, 534–539.
- (56) Wang, J.; Brudvig, G. W.; Batista, V. S.; Moore, P. B. On the relationship between cumulative correlation coefficients and the quality of crystallographic data sets. *Protein Sci.* **2017**, *26*, 2410–2416.
- (57) Wang, J.; Askerka, M.; Brudvig, G. W.; Batista, V. S. Crystallographic data support the carousel mechanism of water supply to the oxygen-evolving complex of photosystem II. *ACS Energy Lett.* **2017**, *2*, 2299–2306.
- (58) Xu, G.; Chance, M. R. Radiolytic modification of acidic amino acid residues in peptides: probes for examining protein-protein interactions. *Anal. Chem.* **2004**, *76*, 1213–1221.
- (59) Kurki-Suonio, K. On the information about deformations of the atoms in X-ray diffraction data. *Acta Crystallogr., Sect. A: Cryst. Phys., Diff., Theor. Gen. Crystallogr.* **1968**, *24*, 379–390.
- (60) Stachowicz, M.; Malinska, M.; Parafiniuk, J.; Wozniak, K. Experimental observation of charge-shift bond in fluorite CaF₂. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2017**, *73*, 643–653.
- (61) Bartholomew, A. K.; Teesdale, J. J.; Hernandez Sanchez, R.; Malbrecht, B. J.; Juda, C. E.; Menard, G.; Bu, W.; Iovan, D. A.; Mikhailine, A. A.; Zheng, S. L.; Sarangi, R.; Wang, S. G.; Chen, Y. S.; Betley, T. A. Exposing the inadequacy of redox formalisms by resolving redox inequivalence within isovalent clusters. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 15836–15841.
- (62) Amin, M.; Askerka, M.; Batista, V. S.; Brudvig, G. W.; Gunner, M. R. X-ray free electron laser radiation damage through the S-state cycle of the oxygen-evolving complex of photosystem II. *J. Phys. Chem. B* **2017**, *121*, 9382–9388.
- (63) Nass, K.; Foucar, L.; Barends, T. R.; Hartmann, E.; Botha, S.; Shoeman, R. L.; Doak, R. B.; Alonso-Mori, R.; Aquila, A.; Bajt, S.; Barty, A.; Bean, R.; Beyerlein, K. R.; Bublitz, M.; Drachmann, N.; Gregersen, J.; Jonsson, H. O.; Kabsch, W.; Kassemeyer, S.; Koglin, J. E.; Krumrey, M.; Mattle, D.; Messerschmidt, M.; Nissen, P.; Reinhard, L.; Sitsel, O.; Sokaras, D.; Williams, G. J.; Hau-Riege, S.; Timneanu, N.; Caleman, C.; Chapman, H. N.; Boutet, S.; Schlichting, I. Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. *J. Synchrotron Radiat.* **2015**, *22*, 225–238.
- (64) Kern, J.; Loll, B.; Luneberg, C.; DiFiore, D.; Biesiadka, J.; Irrgang, K. D.; Zouni, A. Purification, characterisation and crystallisation of photosystem II from *Thermosynechococcus elongatus* cultivated in a new type of photobioreactor. *Biochim. Biophys. Acta, Bioenerg.* **2005**, *1706*, 147–157.
- (65) Fransson, T.; Chatterjee, R.; Fuller, F. D.; Gul, S.; Weninger, C.; Sokaras, D.; Kroll, T.; Alonso-Mori, R.; Bergmann, U.; Kern, J.; Yachandra, V. K.; Yano, J. X-ray emission spectroscopy as in situ diagnostic tool for X-ray crystallography of metalloproteins using an X-ray free-electron laser. *Biochemistry* **2018**, *57*, 4629–4637.
- (66) Sugiura, M.; Inoue, Y. Highly purified thermo-stable oxygen-evolving photosystem II core complex from the thermophilic cyanobacterium *Synechococcus elongatus* having His-tagged CP43. *Plant Cell Physiol.* **1999**, *40*, 1219–1231.
- (67) Cammarata, K. V.; Cheniae, G. M. Studies on 17,24 kD Depleted Photosystem II Membranes: I. Evidences for high and low affinity calcium sites in 17,24 kD depleted PSII membranes from wheat versusspinach. *Plant Physiol.* **1987**, *84*, 587–595.
- (68) Miqyass, M.; van Gorkom, H. J.; Yocum, C. F. The PSII calcium site revisited. *Photosynth. Res.* **2007**, *92*, 275–287.
- (69) Reisman, S.; Ohad, I. Light-dependent degradation of the thylakoid 32 kD protein in isolated chloroplast membranes of *Chlamydomonas Reinhardtii*. *Biochim. Biophys. Acta, Bioenerg.* **1986**, *849*, 51–61.
- (70) Yao, D. C.; Brune, D. C.; Vermaas, W. F. Lifetimes of photosystem I and II proteins in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett.* **2012**, *586*, 169–173.
- (71) Yao, D. C.; Brune, D. C.; Vavilin, D.; Vermaas, W. F. Photosystem II component lifetimes in the cyanobacterium *Synechocystis* sp. strain PCC 6803: small Cab-like proteins stabilize biosynthesis intermediates and affect early steps in chlorophyll synthesis. *J. Biol. Chem.* **2012**, *287*, 682–692.
- (72) Xie, X. S. Biochemistry. Enzyme kinetics, past and present. *Science* **2013**, *342*, 1457–1459.