

## **Supporting Information for:**

Structural Changes of the Oxygen-Evolving Complex of Photosystem II induced by the S<sub>1</sub> to S<sub>2</sub> Transition: A Combined XRD and QM/MM Study

Mikhail Askerka<sup>1</sup>, Jimin Wang<sup>2</sup>, Gary W. Brudvig<sup>\*,1</sup>, Victor S. Batista<sup>\*,1</sup>

<sup>1</sup>*Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107,*

<sup>2</sup>*Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8114, United States*

**Section I. Description of Computational Methods.....2**

**Section II. Electron Density Maps Analysis.....5**

**Section III. References.....8**

## Section I. Description of Computational Methods.

### QM/MM Model Selection

The computational model of the OEC of PSII was constructed as in our previous report:<sup>1</sup> the initial guess for atom coordinates was taken from the 1.9-Å crystal structure (PDB: 3ARC).<sup>2</sup> The model includes residues with C<sub>α</sub> atoms within 15 Å of the atoms in the CaMn<sub>4</sub>O<sub>5</sub> cluster and the two chloride ions near the OEC. Oxygen atoms of water molecules that fall within the 15-Å boundary were also included (85 total). Where the selection caused a gap of up to two residues in a peptide chain, the missing residues were added to provide continuity. Neutral capping groups (ACE/NME) were added for each chain break, with positions determined by the backbone atoms of neighboring residues. A few residues on the periphery of the selection were removed because their side chains extended away from the rest of the selection.

The final protein selection includes the following residues (capping residues in parenthesis use only the backbone atoms):

*D1 (chain A)*: (57)-58-67-(68), (81)-82-91-(92), (107)-108-112-(113), (155)-156-192-(193), (289)-290-298-(299), (323)-324-344:C-terminus

*CP43 (chain C)*: (290)-291-(292), (305)-306-314-(315), (334)-335-337-(338), (341)-342-(343), (350)-351-358-(359), (398)-399-402-(403), (408)-409-413-(414)

*D2 (chain D)*: (311)-312-321-(322), (347)-348-352:C-terminus

Hydrogen atom placements were guessed using the AmberTools12 software package.<sup>3</sup> All acidic residues (ASP, GLU) were modeled as anions. Histidine protonation patterns were determined by inspection: His190 is protonated at N<sub>δ</sub> to be a hydrogen-bond acceptor from D1-Tyr161, His332 is protonated at N<sub>δ</sub> to leave N<sub>ε</sub> as a ligand to Mn2, and His337 is protonated so that N<sub>ε</sub> donates a hydrogen bond to O3. Sodium counter ions were added based on the electrostatic potential outside the protein to neutralize the system. Hydrogen atom placements were refined by 500 steps of minimization using NAMD v.2.8<sup>4</sup> using molecular mechanics force field parameters, with all heavy atoms fixed in their X-ray assigned positions. Charges on the OEC atoms were assigned as previously reported.<sup>5</sup>

### Calculation of the EXAFS

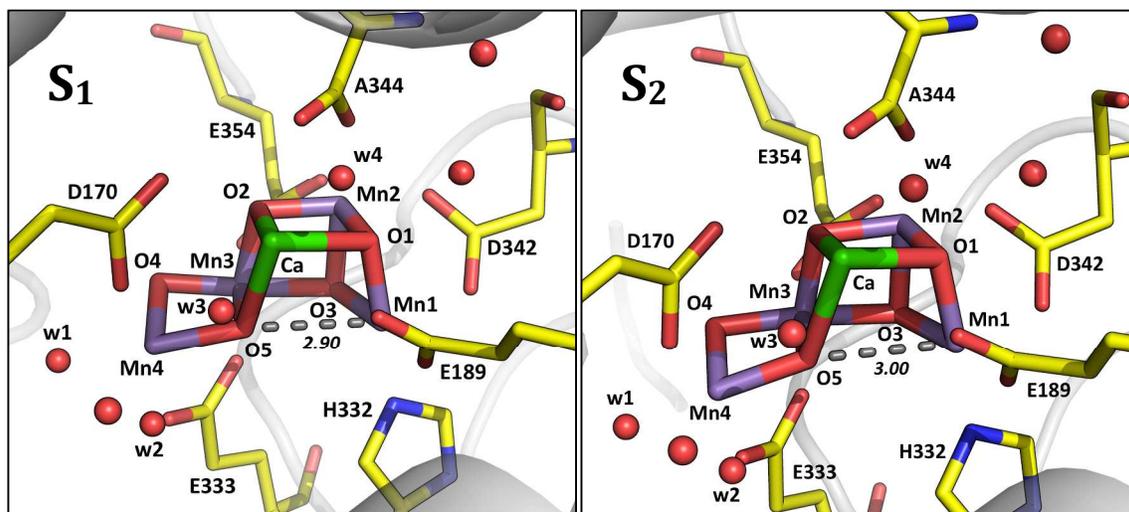
FEFF 8.30<sup>6</sup> combined with IFEFFIT<sup>7</sup> v.1.2.11d programs were used to compute the EXAFS spectra of the S<sub>1</sub> and S<sub>0</sub> models. Only the QM layers were taken into account for computing the EXAFS of the QM/MM optimized structures. For the current calculations, we considered all paths with lengths up to eight scattering legs and the extremely small contribution from hydrogen atoms was not considered. A value of 0.003 Å for the Debye–Waller factors was employed in all calculations. The energy (*E*) axis was converted into photoelectron wave vector (*k*) space by using the following transformation;  $k = (2m_e/(h/2\pi)^2)(E - E_0)$  where *m<sub>e</sub>* is the mass of the electron and *h* is Planck's constant. A value of *E*<sub>0</sub> = 6540.0 eV for the Fermi energy has been employed for the calculations involving the QM/MM model. A fractional cosine-square (Hanning)

window with  $\Delta k = 1$  was applied to the  $k^3$ -weighted EXAFS data. The grid of  $k$  points, which are equally spaced at  $0.05 \text{ \AA}^{-1}$ , was then used for the Fourier transformation (FT) to  $R$  space. A  $k$  range of  $2.29 - 11.5 \text{ \AA}^{-1}$  for the FT for the isotropic EXAFS data was employed. The FT magnitude and EXAFS  $\chi k^3$  values were appropriately scaled to match the experimental data.

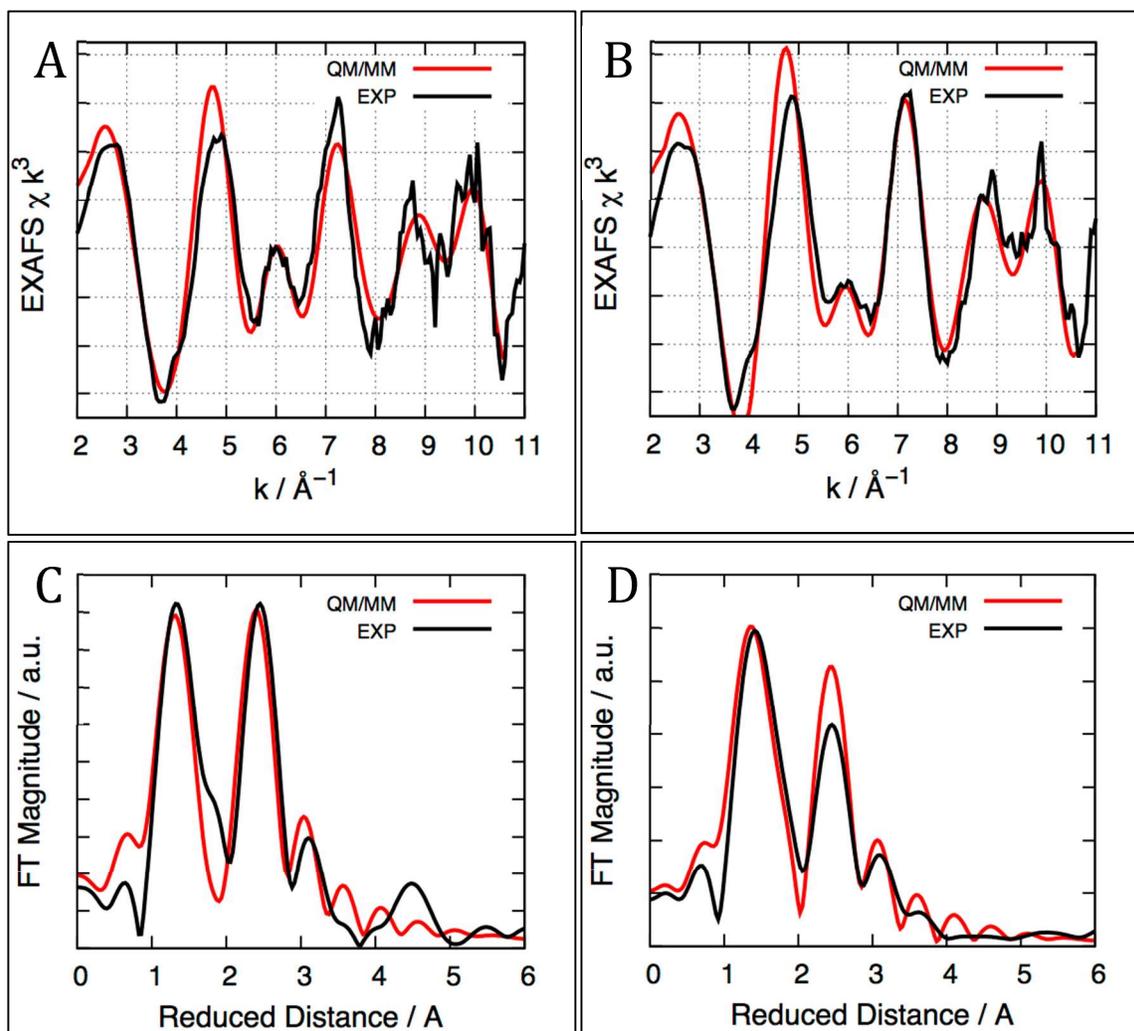
### QM/MM Structure Optimization

The QM/MM optimizations for the  $S_0$  and  $S_1$  states were performed using the ONIOM method<sup>8</sup> as implemented in the Gaussian09 software package.<sup>9</sup> The QM layer was modeled at the B3LYP<sup>10, 11</sup> level of theory using the LaNL2DZ<sup>12, 13</sup> basis set for Mn, Ca and 6-31G\*<sup>14</sup> basis set for C, H, N and O. The QM region was chosen to include the OEC, all directly ligated side chains (D1-D170, D1-E189, D1-H332, D1-E333, D1-D342, and CP43-E354), the C-terminus of D1-A344, hydrogen-bonded residues D1-H337 and CP43-R357, residue D1-D61, and ten water molecules. The anionic ligands were modeled as acetate, D1 C-terminus as propanoate, histidines as methylimidazoles, and arginine as methylguanidium in the QM layer. The AMBER force field<sup>3</sup> was used to model the MM region. All the atoms were allowed to relax during the QM/MM optimizations with the exception of the chloride ion, neutral capping groups (ACE/NME) and oxygen atoms of water molecules in the MM region.

Structures of optimized  $S_1$  and  $S_2$  states as well as the comparison with experimental EXAFS spectra are presented in Figures S1 and S2.



**Figure S1.** The QM/MM-optimized structures of the  $S_1$  and  $S_2$  states of the OEC  $\text{CaMn}_4\text{O}_5$  cluster, along with all directly ligated amino-acid residues (D170, E189, H332, E333, D342, A344 and E354) and 4 directly bonded water molecules (w1 – w4).



**Figure S2.** Comparison between experimental (black)  $k^3$ -weighted EXAFS spectra (A, B) and Fourier transform (FT) magnitudes (C, D) for the OEC of PSII in the  $S_1$  (A, C) and  $S_2$  (B, D) states taken from ref. 15 and the simulated (red) spectra. Reduced space spectra show prominent peaks corresponding to scattering centers in the first (O and N), second (Mn and Ca), and third coordination shells of Mn.

Since our QM/MM studies do not have sufficient constraints on the protein molecule and cannot accurately address any structural changes of the protein molecule during the  $S_1 \rightarrow S_2$  transition, we had to compare the OEC structure between the two states independent of superposition of the protein molecule. We used a least-squares method to superimpose each Mn ion with all 6 ligands at vertices of the Mn coordination octahedron between the two states. This comparison shows that the largest change occurs at Mn4 (0.126  $\text{\AA}$ ) between the two states, whereas Mn1, Mn2, and Mn3 remain relatively unchanged (0.013, 0.020, and 0.074  $\text{\AA}$ , respectively). To determine what type of changes of Mn4 occur during the  $S_1 \rightarrow S_2$  transition, we compared the Mn4 octahedron in each state with an idealized symmetric octahedron and found that Mn4 in the  $S_2$  state is highly

symmetric with a root-mean-squares deviation (rmsd) of 0.060 Å from an idealized octahedron, whereas Mn4 in the S<sub>1</sub> state is more asymmetric with a rmsd of 0.126 Å.

Structures of S<sub>1</sub> and S<sub>2</sub> can be found in PDB format in Supporting Information.

## Section II. Electron Density Maps Analysis.

Our initial calculations were carried out using the 4IXQ and 4IXR coordinates generated by the original authors for the S<sub>1</sub> and S<sub>2</sub> states, respectively.<sup>16</sup> For the following two reasons, we decided to re-calculate the two structures ourselves. First, Kern and colleagues placed the two structures in two different crystallographic origins, making direct comparison nearly impossible. Given the small isomorphous differences between them (22.9%), we chose the same crystallographic origin for both structures. Second, we believe that transferring the high-resolution density maps into the low-resolution structure is more appropriate than transferring atomic coordinates at such low resolution (5.7 Å and 5.9 Å). As reported by Kern and colleagues, the residual  $F_{\text{obs}} - \text{minus} - F_{\text{calc}}$  maps had a large asymmetric distribution of positive features.<sup>16</sup> Some of these features were likely due to missing ordered ligands, lipids, and water molecules in their models, which were partially stripped off during the standard molecular replacement procedure. Our procedure for transferring the electron density maps would keep all associated ligands, lipids, and water molecules as a part of the molecule. This transferring procedure has only 6 parameters of the matrix plus two scaling factors per molecule. Before executing this procedure, we also decided to improve the quality of the maps to be transferred, which required re-refinement of the two additional photosystem II structures with the corresponding PDB accessions of 3ARC and 3BZ1 from two closely related *Thermosynechococcus* genera.<sup>2, 17</sup> During re-refinement, we have corrected several errors in the models, including un-hydrolyzed formyl groups for several initiation methionine residues and modified amino-acid residues embedded inside the membrane. Capped with formyl groups, initiation methionine residues remain uncharged so that they are stable near and inside the membrane. The original authors incorrectly interpreted the formyl groups as multiple conformations of Met side chains.<sup>2</sup> Details of this analysis will be described elsewhere. The re-refinement statistics are: R = 10.5% and R<sub>free</sub> = 17.1% at 1.9 Å for *Thermosynechococcus vulcanus* PSII as opposed to R = 17.5% and R<sub>free</sub> = 20.1% in the original 3ARC model.<sup>2</sup> Continued refinement further improved the R-factors, although this model was not used in this analysis and will be described elsewhere. The refinement statistics are: R = 18.6% and R<sub>free</sub> = 23.7% at 2.9 Å resolution for *Thermosynechococcus elongatus* BP-1 PSII as opposed to R = 24.9% and R<sub>free</sub> = 29.2% in the original 3BZ1 model.<sup>17</sup> We have also identified several small species-specific differences between the two structures, which were ignored for phasing low-resolution structures of the S<sub>1</sub> and S<sub>2</sub> states, in our attempts through multi-crystal non-crystallographic symmetry (NCS) averaging. Since each structure has two copies of PSII in the crystallographic asymmetric unit, the initial model phases for the low-resolution structures of the S<sub>1</sub> and S<sub>2</sub> states were improved using a 6-fold 3-crystal NCS averaging procedure.<sup>18, 19</sup>

To simulate  $F_{\text{obs}}(\text{S}_2) - \text{minus} - F_{\text{obs}}(\text{S}_1)$  difference Fourier maps, we inserted our

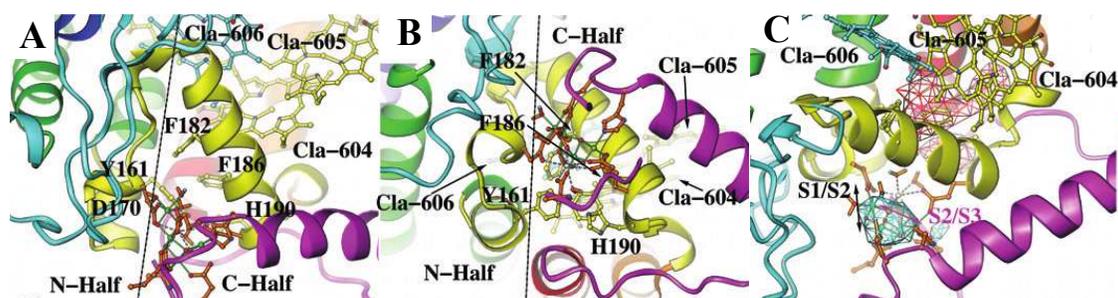
QM/MM  $S_1$ - and  $S_2$ -state models into the corresponding structures of the same crystal for which X-ray data are available.<sup>16</sup> The temperature B-factors for the inserted QM/MM models were replaced by the mean B-factors of the corresponding structures, namely,  $64.35 \text{ \AA}^2$  and  $71.93 \text{ \AA}^2$  for the  $S_1$  and  $S_2$  structures, respectively, according to our re-analysis using the X-ray diffraction data at 5.7- and 5.9- $\text{\AA}$  resolution.<sup>16</sup> Before the insertion of our QM/MM models, the average difference of the calculated phases between the two states was  $12.5^\circ \pm 4.1^\circ$ , and the average difference of calculated amplitudes was 16.0%. With the inserted QM/MM models, the average phase and amplitude differences of calculated structure factors between the two states were  $13.7^\circ \pm 6.1^\circ$  and 16.7%, respectively. We attempted to carry out rigid-body refinement using such hybrid models, but failed to lead to any significant improvements in refinement statistics. This is not surprising since the structural changes observed from our models are indeed very small. The isomorphous difference Fourier method on the other hand is much more sensitive to reveal subtle structural changes associated with any changes in oxidation and electronic state of the OEC (Figure 1).

Features in the difference Fourier maps using the two different phase sets as described were very similar, although the averaged phase set provided slightly cleaner maps and, thus, was used for making the figure reported here. The highest peak next to Mn4 was about  $14.2 \sigma$  when difference densities within the 15- $\text{\AA}$  radius from the center of the OEC were used to calculate the standard deviation. The peak height was to about  $2.9 \sigma$  when difference densities of the entire unit cell were used to calculate the standard deviation. This implies that, although it was the largest feature within 5  $\text{\AA}$  of the OEC next to Mn4, this feature could become undetectable when simulated noise was taken into account using two independent partially refined models, one against the  $S_1$  observed X-ray data, and the other against the  $S_2$  data. That is, the  $F_{\text{calc}}(S_2) - \text{minus} - F_{\text{calc}}(S_1)$  difference maps may not be able to reproduce the  $F_{\text{obs}}(S_2) - \text{minus} - F_{\text{obs}}(S_1)$  features, which were independently confirmed by our simulated  $F_{\text{simulated}}(S_2) - \text{minus} - F_{\text{simulated}}(S_1)$  difference Fourier maps.

The peak in the observed  $F_{\text{obs}}(S_2) - \text{minus} - F_{\text{obs}}(S_1)$  maps on Mn4 in the first subunit was actually higher than the corresponding peak in our simulated  $F_{\text{simulated}}(S_2) - \text{minus} - F_{\text{simulated}}(S_1)$  difference Fourier maps. In the second subunit of the dimeric PSII, the peak appeared at the same location but was reduced to about  $3 \sigma$ , which is close to the level of the simulated maps. From the relative peak heights in our simulations and re-calculated observed difference maps, it is likely that we may have over-estimated the noise level in our simulations since our simulations started with independently partially refined  $S_1$  and  $S_2$  models that may have contained large refinement errors. The reason why the observed difference Fourier maps are so sensitive to very subtle structural changes is that most of refinement errors are cancelled out. The features in the observed difference maps mainly result from the amplitude differences between the observed two states, which are then mapped onto the 3-dimensional space according to the reasonably accurate phase information. We cannot exclude the possibility that additional displacement of the OEC relative to the remaining parts of molecule in the lattice may exist and can contribute to the enhanced features in the observed maps relative to our simulated maps.

In addition to the agreement between the observed  $S_2 - \text{minus} - S_1$  and the simulated

$S_2$ -minus- $S_1$  difference Fourier maps, there is an agreement between the observed  $S_2$ -minus- $S_1$  difference Fourier maps and the residual  $F_{\text{obs}} - \text{minus} - F_{\text{calc}}$  maps in each state as well, where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  denote the observed and calculated amplitudes of the given state, respectively. When the  $S_1$  state was solved using the molecular replacement method starting with the 3ARC model,<sup>2</sup> there was no major difference in the residual  $F_{\text{obs}}(S_1) - \text{minus} - F_{\text{calc}}(S_1)$  maps, indicating that the OEC of the 3ARC model was a good approximation for the  $S_1$  state. When the  $S_2$  state was solved using the 3ARC model, there was a positive feature near Mn4 of the OEC in the residual  $F_{\text{obs}}(S_2) - \text{minus} - F_{\text{calc}}(S_2)$  maps, which coincided with the feature in the simulated and observed difference Fourier maps. This implies that the OEC of the 3ARC model was a poor model for the  $S_2$  state. Thus, the three lines of consistent evidence ascertain that the observed small structural changes near Mn4 are significant.



**Figure S3. Bipartite distributions of subunit A and additional support for the observed  $S_2$ -minus- $S_1$  difference Fourier features.** (A-B) Two views of protein side-chain ligands for the OEC partitioned into the two halves of subunit A with an approximate boundary indicated by a line. (C) The displacements of the metal ions in the OEC during the  $S_1$  to  $S_2$  transition (black double arrows) are approximately orthogonal to the displacements during the  $S_2$  to  $S_3$  transition predicted from our QM/MM studies (magenta double arrows).

### Section III. References.

- (1) Pal, R., Negre, C. F. A., Vogt, L., Pokhrel, R., Ertem, M. Z., Brudvig, G. W., and Batista, V. S. (2013) S<sub>0</sub>-State Model of the Oxygen-Evolving Complex of Photosystem II, *Biochemistry* 52, 7703-7706.
- (2) Umena, Y., Kawakami, K., Shen, J. R., and Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å, *Nature* 473, 55-60.
- (3) D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2012) AMBER 12, University of California, San Francisco.
- (4) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD, *J. Comp. Chem.* 26, 1781-1802.
- (5) Sproviero, E. M., Gascon, J. A., McEvoy, J. P., Brudvig, G. W., and Batista, V. S. (2008) Quantum mechanics/molecular mechanics study of the catalytic cycle of water splitting in photosystem II, *JACS* 130, 3428-3442.
- (6) Ankudinov, A. L., Bouldin, C. E., Rehr, J. J., Sims, J., and Hung, H. (2002) Parallel calculation of electron multiple scattering using Lanczos algorithms, *Phys. Rev. B* 65.
- (7) Newville, M. (2001) IFEFFIT: interactive XAFS analysis and FEFF fitting, *J Synchrotron Radiat.* 8, 322-324.
- (8) Vreven, T., and Morokuma, K. (2000) The ONIOM (our own N-layered integrated molecular orbital plus molecular mechanics) method for the first singlet excited (S<sub>1</sub>) state photoisomerization path of a retinal protonated Schiff base, *J. Chem. Phys.* 113, 2969-2975.
- (9) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr., J. E. P., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. (2009) Gaussian 09, Revision A.1, *Gaussian 09*.
- (10) Becke, A. D. (1988) Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A* 38, 3098-3100.
- (11) Becke, A. D. (1993) Density-functional thermochemistry .3. The role of exact exchange, *J. Chem. Phys.* 98, 5648-5652.

- (12) Hay, P. J., and Wadt, W. R. (1985) Ab initio effective core potentials for molecular calculations - potentials for K to Au including the outermost core orbitals, *J. Chem. Phys.* 82, 299-310.
- (13) Hay, P. J., and Wadt, W. R. (1985) Ab initio effective core potentials for molecular calculations - potentials for the transition-metal atoms Sc to Hg, *J. Chem. Phys.* 82, 270-283.
- (14) Harihara, P. C., and Pople, J. A. (1973) Influence of Polarization Functions on Molecular-Orbital Hydrogenation Energies, *Theor. Chim. Acta* 28, 213-222.
- (15) Haumann, M., Muller, C., Liebisch, P., Iuzzolino, L., Dittmer, J., Grabolle, M., Neisius, T., Meyer-Klaucke, W., and Dau, H. (2005) Structural and Oxidation State Changes of the Photosystem II Manganese Complex in Four Transitions of the Water Oxidation Cycle ( $S_0 \rightarrow S_1$ ,  $S_1 \rightarrow S_2$ ,  $S_2 \rightarrow S_3$ , and  $S_{3,4} \rightarrow S_0$ ) Characterized by X-ray Absorption Spectroscopy at 20 K and Room Temperature, *Biochemistry-US* 44, 1894-1908.
- (16) Kern, J., Alonso-Mori, R., Tran, R., Hattne, J., Gildea, R. J., Echols, N., Glockner, 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., and Yano, J. (2013) Simultaneous Femtosecond X-ray Spectroscopy and Diffraction of Photosystem II at Room Temperature, *Science* 340, 491-495.
- (17) Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A., and Saenger, W. (2009) Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride, *Nat. Struct. Mol. Biol.* 16, 334-342.
- (18) 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., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 67, 235-242.
- (19) Cowtan, K. D. (1994) dmmulti, *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* 31, 34-38.