Supporting Information

Proton exit pathways surrounding the Oxygen Evolving Complex of Photosystem II

Divya Kaur[a,b], Yingying Zhang[b,d], Krystle M. Reiss[c], Manoj Mandal[b], Gary W. Brudvig[c], Victor S. Batista[c], M. R. Gunner[a,b,d,*]

[a]*Department of Chemistry, The Graduate Center, City University of New York, New York, NY 10016, United States*

[b]*Department of Physics, City College of New York, New York 10031, United States*

[c]*Department of Chemistry, Yale University, New Haven, Connecticut 06520, United States*

[d]*Department of Physics, The Graduate Center of the City University of New York, New York, NY 10016, United States*

S1. Figure of PSII highlighting the portion of protein used for MCCE calculations

S2. Water mediated hydrogen bond network for nine representative snapshots

S3. Energetic profile for hydronium moving through the three channels in ten representative snapshots

S4. Details of the PSII molecular dynamics and MCCE simulations setup

S5. Motion of water and chloride molecules observed in MD trajectory

S1. Figure of PSII highlighting the portion of protein used for MCCE calculations

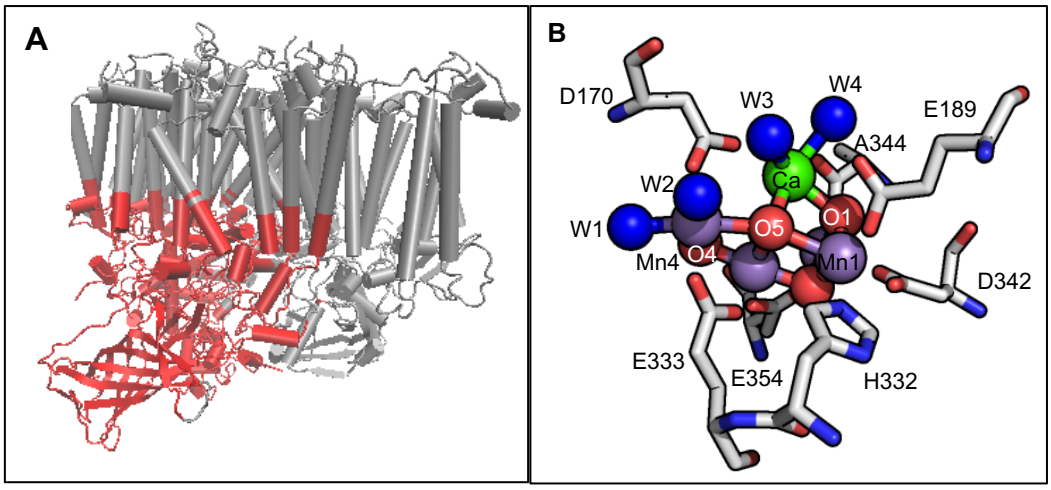
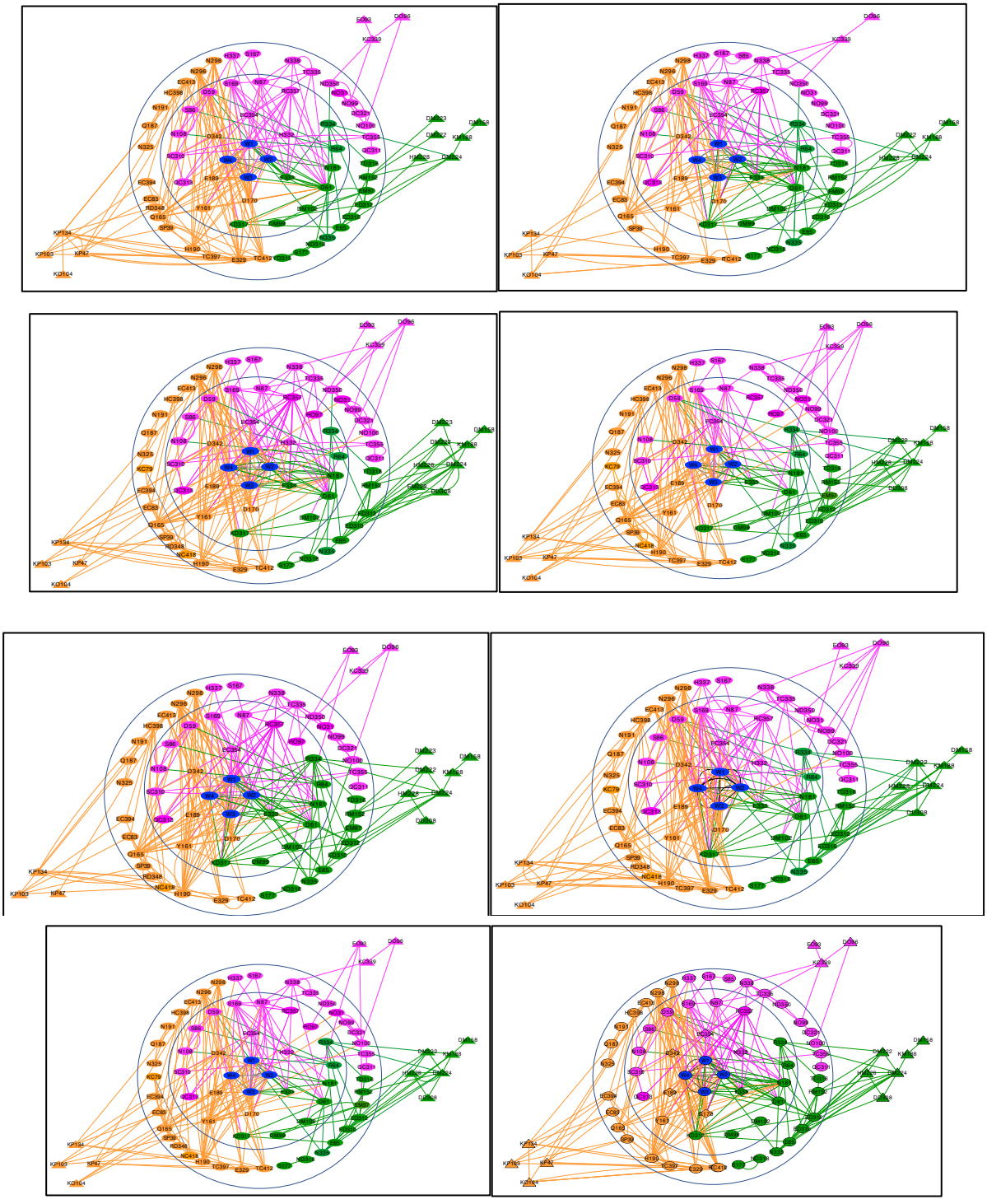


Figure S1 (A) Structure of complete PSII monomer (PDB: 4UB6) [1] in gray color. Hydrogen bond network determined in the red 88 Å × 70 Å × 65 Å rectangle by MCCE. For MCCE, all the atoms in the grey region are deleted. (B) Crystal structure of the OEC (PDB: 4UB6 [4,57]) with all terminal ligands. W1 and W2 are ligands to Mn4 and W3 and W4 to Ca. O5 is nominally closest to the broad channel entry, O4 to the narrow channel and O1 to the large channel.

S2. Water mediated hydrogen bond network for nine representative snapshots



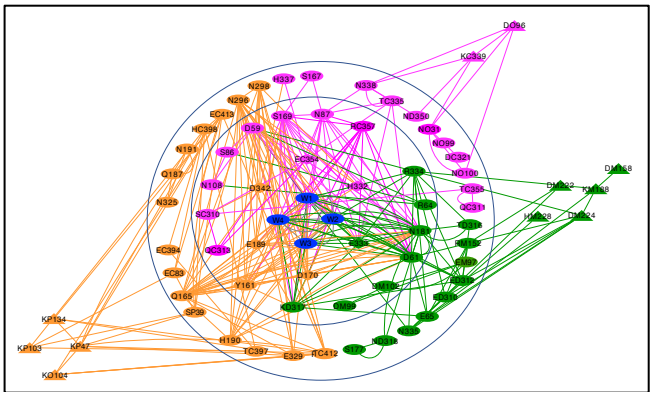


Figure S2 Network of hydrogen bond connections from the OEC to the lumen found in MCCE calculations for nine MD snapshots. Nodes are labeled as Residue type Chain designation Residue number. No chain designation indicates D1. Chain C is CP43; D is D2; M is PsbO; and P is PsbV, O is PsbU. For example, EC354 is CP43-Glu354. Diamonds are primary ligands and triangles are residues with at least 20% of their surface exposed to the lumen. Lines show hydrogen bond connection mediated by 0-4 waters. The inner circle encloses highly interconnected residues near the OEC. Connections between residues nominally in different channels are seen. The outer circle encloses residues in their separated channels. Beyond the outer circle are residues connecting the channel and the surface. The same layout for the residues (nodes) is used in all figures. The changes in lines connecting nodes shows how hydrogen bonds are made and broken in the MD trajectory.

S3. Energetic profile for hydronium moving through the three channels in ten representative snapshots

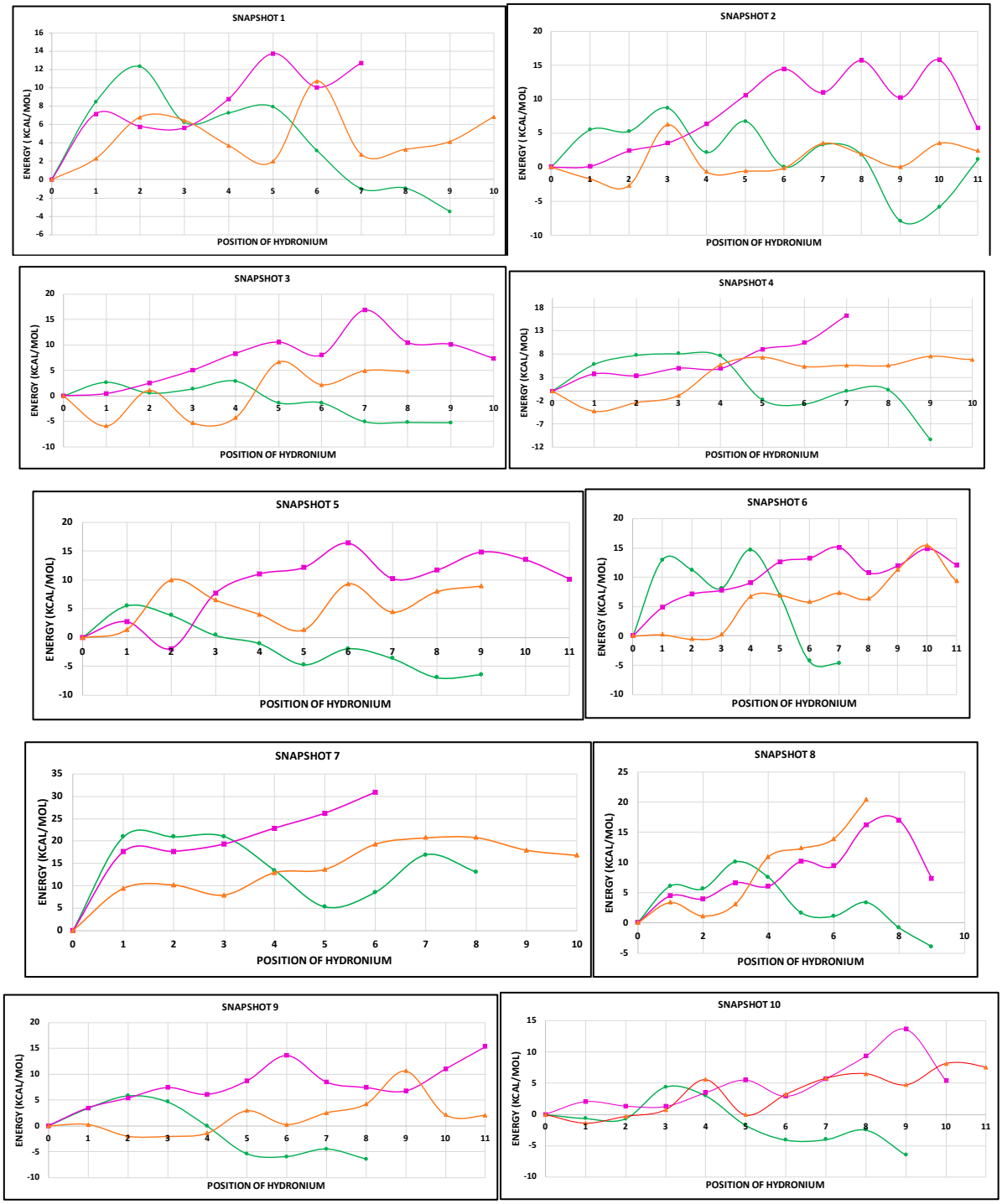


Figure S3 Free energy profile for hydronium at different positions in the broad (green), narrow (magenta) and large (orange) channels for ten individual snapshots. x-axis is the position of hydronium at positions moving away from the OEC in each channel. Residues near each position are described in main text. y-axis is the energy (in kcal/mol) of protein with hydronium at each position. The reference energy (at position 0) is for the full protein with no hydronium plus the energy of an isolated hydronium in water. Shorter lines indicate where a dehydration event breaks the hydrogen bonded chain within the protein, blocking proton egress.

S4. Details of the PSII molecular dynamics and MCCE simulations setup.

MD set up. The MD simulations start with the 1.9 Å resolution X-ray crystal structure of *Thermosynechococcus vulcanus* 4UB6 [1]. The 20 subunits of one PSII monomer are included. The CHARMM-GUI [2] bilayer membrane builder embeds the system within a membrane of MGDG, DGDG and POPG in 1:1:1 ratio. The final system has the PSII monomer in a 180 Å × 180 Å × 81 Å rectangular box with ~300,000 TIP3P [3] water molecules along with the membrane. There are 616 Na^+ and 320 Cl^- ions added to maintain charge neutrality at a 0.15 M concentration. There are 494,990 atoms in the system.

The OEC bond lengths, angles, and dihedrals, as well as those from the Mn to the side chain and water primary ligands, are taken from the QM/MM optimized S_1 state [4]. The cofactor structure is maintained with constraints of 1000 kcal/mol/Å^2 for bond lengths and 200 kcal/mol/Å^2 for angles. Additional restraints of 500 kcal/mol/Å^2 are added for the hydrogen bonds between Q_B and Ser264 and between Y_Z and H190 [5]. The latter may be a strong, short hydrogen bond, which will not be maintained in the standard classical mechanics force field [6]. ESP charges are used for the Ca^{2+} and four Mn (Mn1, Mn2, Mn3 and Mn4). RESP charges are used for the 5 μ-oxo ligands, four terminal waters of Ca^{2+} and Mn4 as well as the primary OEC ligands D1-D170, E189, H332, E333, D342, A344, and CP43-E354. The ESP charges [7] are obtained using density functional theory at B3LYP/6-31G level using the Gaussian software package [8]. RESP charges were derived from the DFT-calculated ESP charges using the RESP program [9]. CHARMM forcefield parameters for lipids and for neutral pheophytin, chlorophyll and plastoquinone cofactors of PSII were taken from Guerra et al. [6].

Ligand bond connections are made for both heme b_{559} and the non-heme Fe complex in CHARMM [10]. The di-sulphide linkage [11] between Cys 19 and 44 in PsbO is made. The protonation states of the residues must be fixed at the beginning of the MD simulations. The ionization states of acidic and basic residues were modified based on MCCE calculations of the protein in the S_1 state with all other cofactors in their ground state [12]. All Asp, Glu, Arg, and Lys are ionized and His, Cys and Tyr are neutral with the following exceptions: CP47 -D380, E387, E405, CP43 -E413, D2 -E242, E343, PsbO- D102, D224, E97, PsbV- K47, K134 are neutral, while His D1- H92, H304, H337, CP47- H343, CP43- H74, H398, PsbO- H228 and PsbU-H81 are protonated. In addition, by default CHARMM-GUI chooses the neutral His with a proton on ND1 (HSD). However, MCCE found that the following His residues prefer to have a proton on NE2 (HSE): D1- H195, H252, CP43- H157, H201, D2- H61, H87, H189, H336, PsbO- H231, PsbV- H118. It should be noted that these choices have consequences. For example, choosing the HSD tautomer for D1- His252 leads to the opening of a loop near Q_B.

OpenMM [13] is used to generate MD trajectories including periodic boundary conditions with Langevin dynamics with Nose-Hoover Langevin piston at constant pressure (1 bar) and temperature

(303.15 K). The coulombic interactions are calculated using the Particle-Mesh Ewald (PME) algorithm. A 2fs time step is used. To relax initials clashes, the system is first minimized in CHARMM using Steepest Descent (SD) minimization for 1000 steps. The system is then further equilibrated in OpenMM where the protein backbone, side chains and lipids are allowed to relax for 355 ps when all constraints are removed. The production run is 100 ns long.

The position of chloride near the OEC has moved ≈3-5 Å in the trajectory from its position within the structure. This is a rigid body motion, retaining stable distances amongst Cl and D1-E333, R334 and D2-K317. The Cl to D1-N181 lengthens in the new minimum. The calculations were restarted and equilibrated several times and this motion was found each time.

MCCE set up. The implicit salt concentration is 0.15 M with a 2 Å stern layer, and a pH of 6. The dielectric constant is 4 for the protein and 80 for the solvent. Parse charges [14], optimized for continuum calculations, are used for amino acids while integer valence charges are used for the OEC [12]. TIPS [3] charges are used for explicit waters. Standard MCCE topology files are used for the cofactor chlorophylls, pheophytin, plastoquinone, chloride, heme and non-heme iron [15]. Amber van der Waal parameters are reduced to 25 % of their full value based on previous MCCE benchmark calculations [16]. As we are concerned with the proton exit from the OEC which is out of the intra-membrane region no lipids are included. The coordinates for the ten snapshots, including only those atoms in the region used for MCCE analysis can be found in the GitHub repository (https://github.com/DivyaMatta15/Proton-egress-pathways-surrounding-oxygen-evolving-complex-of-Photosystem-II).

S5. Motion of water and chloride molecules observed in the MD trajectory

The water molecules in the trajectories are allowed to leave or come back near the water channels. Notably, near the narrow channel a chain of water molecules is aligned from the OEC extending from the O4-μ-oxo bridge to D1-N338. This was also observed in previous computational studies that focused on the narrow channel [7,17]. Recently, computational studies done on these trajectories by Batista et al find that the first 4-5 waters in narrow channel have a very clear hydrogen bond network, with well positioned water positions. The oxygen positional probability distributions are spherical and visible even at high contour levels, indicating they are maintained in these positions. The orientation of the water proton positions is also maintained through the trajectory.

Reference

[1] Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, Crystal structure of oxygen-evolving photosystem II at a resolution of 19 Å, Nature, 473 (2011) 55–60.

[2] J. Lee, X. Cheng, J.M. Swails, M.S. Yeom, P.K. Eastman, J.A. Lemkul, S. Wei, J. Buckner, J.C. Jeong, Y. Qi, S. Jo, V.S. Pande, D.A. Case, C.L. Brooks, A.D. MacKerell, J.B. Klauda, W. Im, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field, J Chem Theory Comput, 12 (2016) 405–413.

[3] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys., 79 (1983) 926–935.

[4] S. Luber, I. Rivalta, Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, G.W. Brudvig, V.S. Batista, S_1-state model of the O_2-evolving complex of photosystem II, Biochemistry, 50 (2011) 6308–6311.

[5] H. Ishikita, K. Saito, Proton transfer reactions and hydrogen-bond networks in protein environments, J R Soc Interface, 11 (2014) 20130518.

[6] F. Guerra, S. Adam, A.-N. Bondar, Revised force-field parameters for chlorophyll-a, pheophytin-a and plastoquinone-9, J. Mol. Graph. Model., 58 (2015) 30–39.

[7] K. Reiss, U.N. Morzan, A.T. Grigas, V.S. Batista, Water network dynamics next to the oxygen-evolving complex of photosystem II, Inorganics, 7 (2019) 39.

[8] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloini, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J.V. Ortiz, A.F. Izmaylov, et al., G09 | Gaussian.com, (2016).

[9] D.A. Case, T. Darden A., T.E. Cheathem, 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, K.F. Kolossvai, K.F. Wong, F. Paesani, J. Vanicek, et al., AMBER 12, University of California, San Francisco, (2012).

[10] S. Adam, M. Knapp-Mohammady, J. Yi, A.-N. Bondar, Revised CHARMM force field parameters for iron-containing cofactors of photosystem II, J Comput Chem, 39 (2018) 7–20.

[11] J. Nikitina, T. Shutova, B. Melnik, S. Chernyshov, V. Marchenkov, G. Semisotnov, V. Klimov, G. Samuelsson, Importance of a single disulfide bond for the PsbO protein of photosystem II: protein structure stability and soluble overexpression in Escherichia coli, Photosyn. Res., 98 (2008) 391–403.

[12] D. Kaur, W. Szejgis, J. Mao, M. Amin, K.M. Reiss, M. Askerka, X. Cai, U. Khaniya, Y. Zhang, G.W. Brudvig, V.S. Batista, M.R. Gunner, Relative stability of the S_2 isomers of the oxygen evolving complex of photosystem II, Photosyn. Res., 141 (2019) 331–341.

[13] P. Eastman, J. Swails, J.D. Chodera, R.T. McGibbon, Y. Zhao, K.A. Beauchamp, L.-P. Wang, A.C. Simmonett, M.P. Harrigan, C.D. Stern, R.P. Wiewiora, B.R. Brooks, V.S. Pande, OpenMM 7: Rapid development of high performance algorithms for molecular dynamics, PLoS Comput. Biol., 13 (2017) e1005659.

[14] D.J. Tannor, B. Marten, R. Murphy, R.A. Friesner, D. Sitkoff, A. Nicholls, B. Honig, M. Ringnalda, W.A. Goddard, Accurate first principles calculation of molecular charge distributions and solvation energies from ab initio quantum mechanics and continuum dielectric theory, J. Am. Chem. Soc., 116 (1994) 11875–11882.

[15] M. Amin, L. Vogt, W. Szejgis, S. Vassiliev, G.W. Brudvig, D. Bruce, M.R. Gunner, Proton-coupled electron transfer during the S-state transitions of the oxygen-evolving complex of photosystem II, J. Phys. Chem. B, 119 (2015) 7366–7377.

[16] M.R. Gunner, X. Zhu, M.C. Klein, MCCE analysis of the pK_as of introduced buried acids and bases in *Staphylococcal nuclease*, Proteins, 79 (2011) 3306–3319.

[17] K. Saito, A.W. Rutherford, H. Ishikita, Energetics of proton release on the first oxidation step in the water-oxidizing enzyme, Nat Commun, 6 (2015) 8488.