BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **XXXX,** *XXX,* 000–000 A
DOI: 10.1021/bi901900s

Tracking the Flow of Water through Photosystem II Using Molecular Dynamics and Streamline Tracing[†]

Serguei Vassiliev,* Pascal Comte, Abdullah Mahboob, and Doug Bruce

Department of Biology, Brock University, 500 Glenridge Avenue, St. Catharines, Ontario L2S 3A1, Canada

Received November 5, 2009; Revised Manuscript Received January 29, 2010

ABSTRACT: The $CaMn_4$ cluster of the oxygen-evolving complex (OEC) of photosynthesis catalyzes the light-driven splitting of water into molecular oxygen, protons, and electrons. The OEC is buried within photosystem II (PSII), a multisubunit integral membrane protein complex, and water must find its way to the $CaMn_4$ cluster by moving through protein. Channels for water entrance, and proton and oxygen exit, have previously been proposed following the analysis of cavities found within X-ray structures of PSII. However, these analyses do not account for the dynamic motion of proteins and cannot track the movement of water within PSII. To study water dynamics in PSII, we performed molecular dynamics simulations and developed a novel approach for the visualization of water diffusion within protein based on a streamline tracing algorithm used in fluid dynamics and diffusion tensor imaging. We identified a system of branching pathways of water diffusion in PSII leading to the OEC that connect to a number of distinct entrance points on the lumenal surface. We observed transient changes in the connections between channels and entrance points that served to moderate both the flow of water near the OEC and the exchange of water inside and outside of the protein. Water flow was significantly altered in simulations lacking the OEC which were characterized by a simpler and wider channel with only two openings, consistent with the creation of an ion channel that allows entry of Mn^{2+}, Ca^{2+}, and Cl^- as required for construction of the $CaMn_4$ cluster.

Photosystem II (PSII)[1] is a pigment−protein complex embedded in the thylakoid membranes of plant chloroplasts and cyanobacteria that catalyzes the light-induced oxidation of water and reduction of plastoquinone in oxygenic photosynthesis (*1*, *2*). This catalytic function is performed by light-driven electron transfer reactions through redox cofactors of the PSII reaction center. By coupling water oxidation to photochemistry, PSII taps into an abundant energy supply that ultimately powers most forms of life. The heart of PSII's unique ability is the oxygen-evolving complex (OEC), consisting of a cluster of four high-valence manganese ions and a calcium ion. Water cleavage requires accumulation of four oxidation equivalents in the $CaMn_4$ cluster according to the S-cycle proposed by Joliot and Kok (*3*, *4*).

The oxidation of water to O_2 requires an extremely strong oxidant and is prone to side reactions resulting in the formation of reactive oxygen species. The OEC is surrounded by a hydrophobic pocket and buried deep in the protein to prevent excessive access of water and solute to the catalytic site (*5*). The restricted access of large competitive water analogues to the water splitting reaction first led to the proposal that the site of water oxidation was protected by the protein (*6*). Subsequent studies confirmed this idea by showing that perturbation of the protein, via release of extrinsic PSII polypeptides, resulted in the formation of H_2O_2 at the expense of O_2 evolution, a condition reversed by the addition of kosmotropic solutes that promote protein aggregation (*5*). Access to the OEC was probed in an ESEEM study that showed the level of binding of alcohols to the $CaMn_4$ cluster decreased with the increasing size of the alcohol molecule, leading the authors to suggest the presence of a water channel (*7*). As O_2 and protons are released from the same catalytic site, channels for the exit of both have also been suggested (*8*, *9*). With requirements for three potentially independent channels, the "plumbing" of PSII appears to be complex. Additional complexity arises during the assembly or repair of PSII. Insertion of the $CaMn_4$ cluster occurs after protein assembly and thus requires the presence of an ion channel to facilitate the movement of Mn^{2+}, Ca^{2+}, and Cl^- ions through the protein to the site of the OEC.

A search for channels became possible after publication of crystal structures of the PSII core complex (*10*, *11*). The "Caver" program was used to search for connecting cavities that might form channels within the structure of Protein Data Bank (PDB) entry 1S5L (*12*). Three channels were found. The widest and the least hydrophilic (i) was suggested to be for oxygen. A more polar channel (ii) was assigned to water and protons, and the most polar channel (iii) was designated for protons alone. Note that the Caver algorithm searches for the shortest pathways of a chosen minimal diameter from a given point inside the protein to the surface. The application of this program thus biases the search for channels toward a set of round "pipelines". Later, a comprehensive identification of all cavities near the OEC in the structure of PDB entry 2AXT, regardless of whether they connected to the surface, revealed a complex system of potential channels (*13*). The approach taken in that work was a step forward in understanding how water is carried to the active site as it revealed a complex interconnected network of water pores and cavities inside the PSII protein. The most recent structure of PSII

[†]This work was supported by Discovery and Equipment grants from the Natural Science and Engineering Research Council of Canada.

*To whom correspondence should be addressed. E-mail: svassili@brocku.ca. Telephone: (905) 688-5550. Fax: (905) 688-1855.

[1]Abbreviations: PSII, photosystem II; RC, reaction center; OEC, oxygen-evolving complex; CTD, protein C-terminal domains; MD, molecular dynamics; QM, quantum mechanics.

(PDB entry 3bz1) was again analyzed with Caver (*14*, *15*) which generated a system of eight channels originating at the OEC that merged into six exits on the lumenal side of PSII. The largest channels were in positions similar to those of the "water channels" found in earlier studies. A number of "new" channels with smaller minimum diameters were also identified, and on the basis of their size and the placement of hydrophilic residues, they were assigned as proton channels.

To date, efforts to identify water channels within PSII have been based on the analysis of cavities within X-ray structures. There are two problems with this approach.

(1) The existence of a cavity in a crystallographic structure does not necessarily mean that it is occupied by water. Placing explicit water molecules and assessing their potential is the way to solve this problem. Such an approach was used in a DFT-QM/MM study of OEC which identified a chain of hydrogen-bonded water molecules approaching the $CaMn_4$ cluster along two distinct pathways (*16*, *17*). This study, however, was limited by the inclusion of only a small protein area adjacent to the active site of water splitting.

(2) Channels in static structures have fixed boundaries and a rigid "brick wall" lining (*12*, *13*). This is an unrealistic picture as proteins and the channels within them are dynamic at physiological temperatures. Thermal motions of the protein will continuously modify channel networks and may lead to transient opening and closing of some pores (*18*). These processes may be critical in controlling the movement of water with the protein. As proposed in ref *19*, dynamic simulations of water diffusion are essential to the understanding of water channels and water movement in PSII.

Dynamic simulations of water diffusion pathways have been limited to proteins in which well-defined channels with relatively simple geometry have been identified, such as aquaporins (*20*−*23*) and lysozyme crystals (*24*, *25*). In cases like these, standard simulation analysis methods, including radial distribution functions, mean square displacement, and calculation of solvent density along the direction of the channel, are sufficient. However, large protein systems, like PSII, often have multiple nanopores with complex shapes, different sizes, interconnections, and branch points. Dynamics of water may be further complicated by formation of transient nanopores (*18*). For such systems, identification of the most probable pathways from the active site to the bulk solvent becomes a major problem.

Unlike previous static cavity searches, in this work we have identified energetically favorable locations for water molecules in all cavities of PSII protein and placed explicit water molecules within them. We then simulated motion of the whole PSII core complex, including water at physiological temperature, and, finally, developed a novel way of looking at the movement of water into and through proteins. Our streamline tracing approach was used to map the patterns of equilibrium water flow within PSII. We identified a network of dynamic water channels leading to and controlling the flow of water to the OEC. In addition, our simulation of PSII in the absence of the OEC revealed the formation of a shorter and wider channel consistent with the movement of metal ions into the active site required for assembly of the $CaMn_4$ cluster.

METHODS

Molecular Dynamics Simulations. Our MD simulations were similar to that described in (*26*) with the following changes. The current simulation was based on the 2AXT X-ray structure of the PSII core complex with the unassigned subunit X removed. Water was added using DOWSER (*27*). The protocol used by DOWSER was to place water in internal cavities, to relax them to minimize their interaction energy, and finally to select and retain only water molecules with energies of less than −12 kcal/mol (for further details, see ref *27*). The MD simulations were conducted with NAMD (*28*) and performed with periodic boundary conditions at a constant pressure (1 atm) and temperature (300 K) using Langevin dynamics with Nosé-Hoover Langevin piston pressure control. The long-range electrostatic interactions were calculated using the particle mesh Ewald algorithm. Both the electrostatic and the Lennard-Jones interactions had a twin-range cutoff of 10−12 Å. The AMBER-1999 force field (*29*) with modification of the ϕ and ψ dihedral parameters as described in ref *30* was used for protein, and the GLYCAM-2000a force field (*31*) was used for the headgroups of galactolipids. The force field for cofactors was based on the parameters described previously (*26*). The geometry and partial atomic charges of the OEC were taken from the QM-refined model of the OEC and its ligands in the S1 state described in ref *32*. After an initial equilibration period of 2 ns during which the protein backbone and heavy atoms of all cofactors were constrained to their initial X-ray positions and side chains were allowed to relax, all constraints were removed and a 10 ns production run was performed. The whole system, including the PSII complex, the membrane patch, and TIP3P water, contained 320000 atoms; its size was 183 Å × 132 Å × 132 Å.

Fiber Tracking Technique. Fiber tracking is based on the determination of the velocity field and subsequent calculation of the motion of the particle in it. We applied this approach to track the movement of individual water molecules within our larger MD simulation of PSII. To implement this technique, time-averaged three-dimensional direction field data must be extracted from the dynamics simulation by determination of a diffusion tensor. To calculate the tensor field, we divide the simulation system volume into n small cubic elements (voxels) with a volume of ~1 $Å^3$ each. For all voxels containing a water molecule at time t, we find the position of this water molecule at time $(t + \Delta t)$ and calculate the tensor elements according to the Einstein relation:

$$T^{\alpha\beta} = \frac{\langle \alpha(t+\Delta t) - \alpha(t) \rangle \langle \beta(t+\Delta t) - \beta(t) \rangle}{2\Delta t} \quad \alpha, \beta = \{x, y, z\}$$

To obtain all nine tensor elements, we sequentially substituted α and β with x, y, and z. Tensor elements are averaged over the time window of the MD run, and the diffusion tensor is diagonalized to find the direction of fastest motion:

$$\mathbf{T} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}^{\mathrm{T}}$$

Eigenvalue λ_i of the diffusion tensor describes the diffusion rate along the direction of the corresponding eigenvector $\mathbf{v}_i$. Channels (fibers) are characterized by anisotropic linear diffusion ($\lambda_1 > \lambda_2 \cong \lambda_3$). The eigenvector associated with the largest eigenvalue of the diffusion tensor is parallel to the local channel direction (*33*). Therefore, the vector tangent to the trajectory at point s is equal to the eigenvector calculated at this point, and the channel trajectory is described by a system of three differential equations:

$$\frac{d\mathbf{r}}{ds} = \mathbf{v}_1[\mathbf{r}(s)]$$

In this equation, $\mathbf{r}$ is a three-dimensional vector defining a position in space, $\mathbf{v}_1$ is the principal eigenvector, obtained as

described in the previous paragraph, and s is a parameter of the parametric curve that has values $s(x,y,z)$. To track a channel, we solve the system numerically for an initial condition which specifies a starting point: $\mathbf{r}(0) = \mathbf{r}_0$. Solution of this system of differential equations requires a continuous and smooth direction field. Due to a limited time frame in molecular dynamics simulations, the diffusion tensor data are discrete, noisy, and coarsely sampled.

To perform reliable and robust fiber tracking, we generated a continuous smooth representation of the direction field using a trilinear interpolation on the tensor matrix elements (*33*) and implemented a moving least-squares filter that approximates data locally with a low-degree polynomial considering the location, orientation, and history of motion (*34*). To find all channels originating in a region of interest, we first identified "seed points", voxels with high linear anisotropy: $c(x,y,z) = (\lambda_1 - \lambda_2)/(\lambda_1 + \lambda_2 + \lambda_3)$. We then integrated the streamline from each of the seed points in both directions. We did not allow acute turns, and we stopped tracing when the angle between the previous and current directions exceeded 80°. We also stopped tracing when the length of the streamline exceeded the maximum dimension of the tensor field. When deciding whether to accept the next point or to stop tracing, we also checked if either the anisotropy or diffusion rate dropped below allowed threshold values. Typical values for these thresholds in our study, respectively, were as follows: $c > 0.2$ (meaning that fastest motion is at least 1.5 times faster than motion in any other directions), and $\lambda > 0.01$ Å^2/ps. For comparison, the experimental diffusion rate of water is 0.25 Å^2/ps. Breaks in streamlines will occur in areas of diffusion tensor singularities where the diffusion tensor is a perfect disk, or a perfect sphere. When tracking algorithms encounter such a singularity, the principal direction will be undefined. Such a break may occur even though water movement continues. For example, a singularity can occur when streamlines enter a wide cavity where a strong directionality of motion is lost or at junction points where multiple streamlines moving in different directions meet. Such areas can be differentiated from physical blocks to water flow by inspection of the structure in the region of the break.

RESULTS

MD Simulations of Fully Solvated PSII with and without the OEC. We performed MD simulations of PSII in both the presence and absence of the OEC. These simulations allowed us to investigate mechanisms responsible for controlling the access of water to the $CaMn_4$ cluster and to look for channels potentially responsible for delivering metal ions to the active site.

To solvate PSII within our MD simulations, we placed water molecules using the Dowser program (see Methods). More than 900 buried water molecules were added to each of the PSII simulations (Figure 1A). Most of the internal water molecules were located in the area of the "large channel system" as described in ref *13*. Initial water placement suggested multiple entrance points and pathways to the OEC.

Each MD system was simulated for 10 ns after an initial 2 ns equilibration. During the simulations, we observed the equilibrium motion of water across the protein−water interface at multiple points as well as water motion within the protein (Figure 1B,C). Trajectories of motion of several representative water molecules for our simulation in the presence of the OEC are shown in Figure 1B to illustrate the different mobilities of water molecules in different regions of PSII. Identification of time-averaged structures for all of the nanopores inside PSII would be very difficult in this large and complex system. Hence, we developed a specialized data analysis technique to visualize water movement within our PSII MD simulations.

Tracking Water Streams inside PSII. To map water movement within PSII, we designed a tool to locate areas characterized by highly anisotropic motions of water, "streams". We implemented a streamlining technique, "fiber tracking", originating from fluid dynamics. This technique is used in MRI to measure the location, orientation, and anisotropy of water tracts in soft fibrous tissues (*33*, *35*−*38*). Our fiber tracking approach depends upon the extraction of three-dimensional velocity information for individual water molecules within our MD simulation of PSII. We use the water velocity information to construct "streamlines" that show the most probable pathways of water diffusion in the protein.

Streamline analysis differs from following the motion of individual water molecules in several aspects. Individual water molecules in microscopic systems undergo stochastic Brownian motion as illustrated in Figure 1B,C,E. In contrast, individual fibers in a streamline analysis are calculated from the overall diffusional motion of all water molecules found in each volume element (voxel) of the simulated system over the entire time course of the simulation. This analysis can be performed if time- and ensemble-averaged dynamics of the water molecules are time-independent, which is the case when a molecular dynamics system reaches a stationary equilibrated state. Fiber tracks are started in voxels that exhibit the highest anisotropy of water diffusion. Neighboring voxels with a high diffusion rate favoring a particular direction are then connected as described in Methods by tracking algorithms to give streamlines. Streamlines, thus, represent areas where water molecules have moved the most in a directional manner during the course of the MD simulation. The time and ensemble averaging implemented in streamline tracing makes the technique exceptionally robust and capable of detecting streams connecting locations even if none of the water molecules have traveled the whole path during the simulation time window.

When applied to our PSII MD simulation in the presence of the OEC, our fiber tracking technique found streamlines forming a complex network leading from the surface of the protein to the buried $CaMn_4$ cluster. Water streamlines were found in spaces free of protein backbone atoms and in most cases free of all protein heavy atoms. Some streams were found in areas transiently occupied by protein side chains. In general, we found the streamline system to be located in the same regions as the cavities found previously in analyses of the static structure (*12*−*14*). Motions of individual water molecules matched the streamline direction but also exhibited some random deviations (Figure 1E), which is expected due to the stochastic nature of diffusion as described above.

Comparison of Water Streams with Previously Discovered Channels. Water streamlines and entrance points obtained from our analysis of an unconstrained MD simulation of the PSII core complex, including the OEC, are shown in Figure 2. We found water streams in all four of the channels or channel systems described in ref *13*, which had been named "broad", "back", "narrow", and "large".

The back channel of ref *13* was originally identified in ref *12* as an oxygen channel and called channel (i). In ref *13*, the authors proposed that two connected channels, (i) and (ii) of ref *12*,

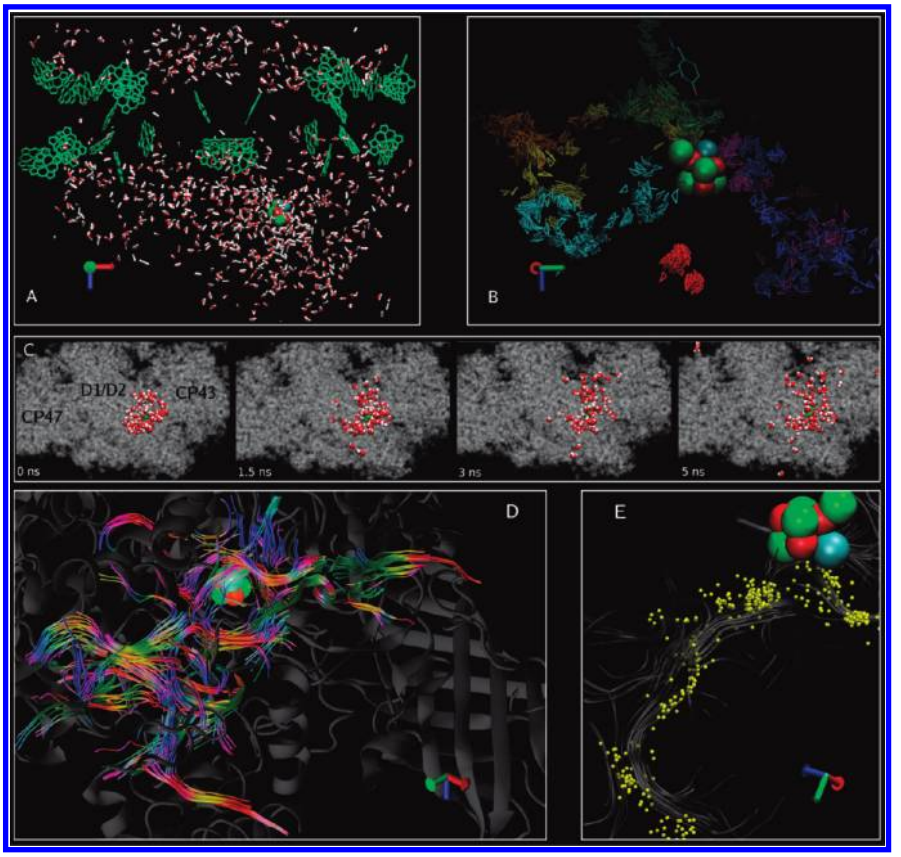

FIGURE 1: Structure and dynamics of water in the MD simulation of PSII in the presence of the OEC. The *x*, *y*, and *z* (red, green, and blue, respectively) axes define the world coordinate system which stays the same in all subsequent figures: the *z* axis is normal to the plane of the thylakoid membrane. The Mn, Ca, and O atoms of the OEC are depicted as green, blue, and red spheres, respectively. (A) Water molecules in PSII placed using Dowser (see the text). Core antenna and reaction center chlorins are colored green. (B) Trajectories of motion of several representative water molecules (each in a different color) inside PSII. Tyrosine Z is shown in licorice format. (C) Snapshots from the dynamics simulation showing diffusion of water molecules inside the protein. Water molecules within 10 Å of the OEC were selected (0 ns) and viewed again at later time intervals (1.5, 3, and 5 ns). Protein (gray) has been clipped for the sake of clarity at the level of the OEC. This view is along the *z* axis looking from the lumen. Core PSII antenna (CP43 and CP47) and reaction center (D1 and D2) proteins are labeled. (D) Water streamlines in the interior of PSII obtained from analysis of the 10 ns MD simulation (see the text for details). The protein has been clipped for the sake of clarity. Streamlines are color coded by direction: red streamlines propagate along the *x* direction, green *y*, and blue *z*. The protein is shown as a gray ribbon. (E) Streamlines (gray) overlapped with the trajectory of one representative water molecule from the MD run (yellow).

should be considered together and that both branches could supply substrate water to the $CaMn_4$ cluster, as well as allow the exit of molecular O_2. Our streamline tracing analysis (Figure 3A) revealed a number of water streams in this area with an interruption of continuity in a hydrophobic area of the back channel (near CP43-Phe358, CP43-Phe292, and D1-Leu91) where vortex streams (rotating around CP43-Pro307 and Asp360) appear before the channel makes an almost 90° turn toward the OEC. The recent Caver analysis of the 2.9 Å structure (*14*) indicates this region as a branching point between their A1 and A2 channels. Our streamlines and the back channel or "A2" channel exit the protein between D1-His92 and CP43-Asp360 (exit 1 in Figure 2; the residues are shown in Figure 3A).

Previously, the broad channel, as described in ref *13*, was found to be blocked ~12 Å from the OEC and did not extend all the way to the surface. The broad channel overlapped with the proximal part of channel (iii) as identified in ref *12*. Beyond this point, the broad channel was blocked by four residues (D1-Glu65, -Pro66, and -Val67 and D2-Glu312). This channel was identified as channel "C" in ref *14*. In all previous structural studies, this channel has been proposed to be exclusively for proton exit because of its narrow diameter and hydrophilic residues (*12*−*14*). In contrast to this conclusion, we found water streams in the broad channel (Figure 3B) going around the blockage, extending to the surface, and forming two entrance points: near PsbO-Pro201, -Ala203, and -Glu205 and CP47-Arg422 and -Lys423 and near D1-Glu65 and -Arg334 (exits 2 and 3 in Figure 2). After the blockage, our streamlines followed a path different from that of channel (iii), and we have not found any streams in the distal part of channel (iii).

The narrow channel was not reported in ref *12* and was first described in ref *13*. The narrow channel was open all the way to the surface, exiting at CP47-Arg385, -Ala386, and -Ser388 and PsbU-Leu47, -Tyr51, -Leu121, -Asp126, and -Asn130 (exit 4 in Figure 2), and was assigned as an alternative H^+ exit pathway (*13*). This channel was identified as channel "F" in ref *14* and also designated for protons. The latest study (*15*) clarified that channels D−F overlap substantially with each other and with the narrow channel of ref *13*. We found water streams in the area of the narrow channel (Figure 3C). Streams are interrupted approximately halfway along this channel in the area of a T-junction between streamlines. Most likely, this streamline interruption arises from low anisotropy at the junction point and does not mean that water does not move across the junction area. One of the streams (the stream going up from the junction in Figure 3C) leads out of the narrow channel area and connects it to the large channel system (Figure 3D). Another stream connects the narrow channel to the broad channel (Figure 3E).

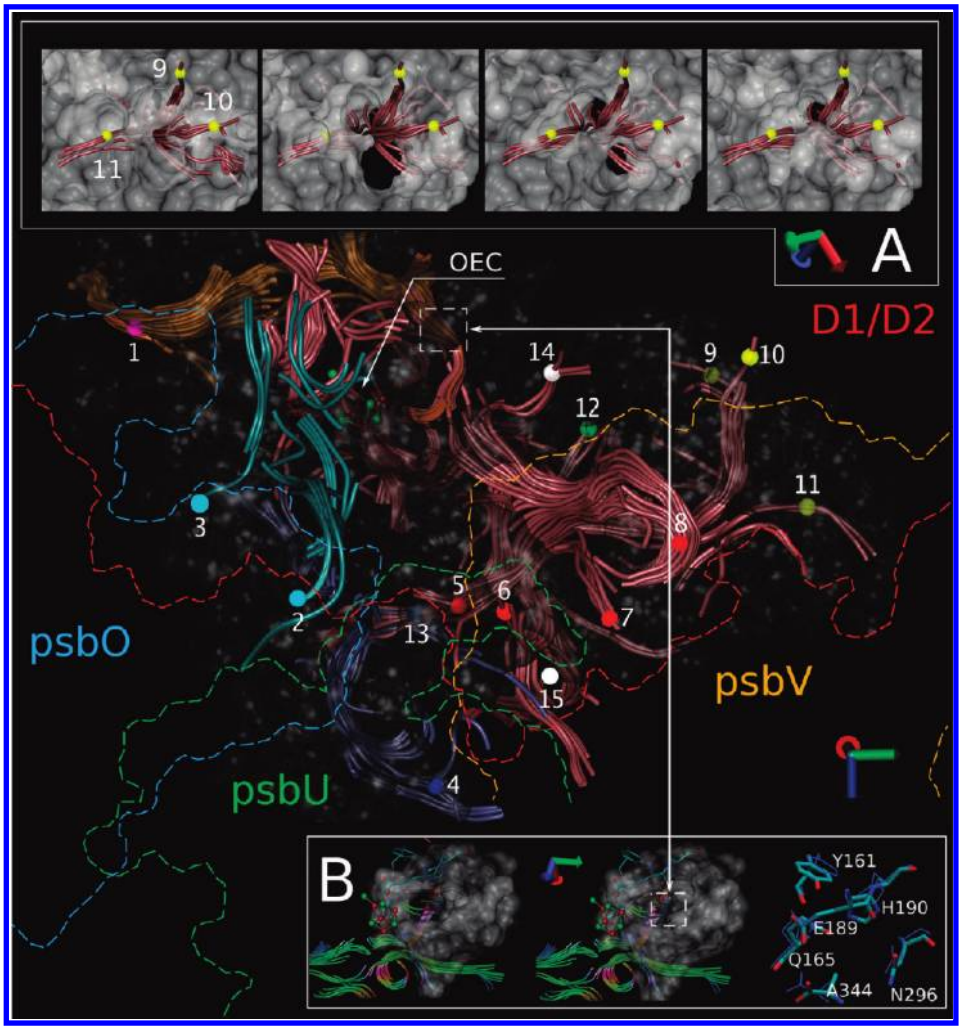

FIGURE 2: Water streamlines in PSII in the presence of the OEC. The surface of the protein immediately surrounding the previously described broad, narrow, and back channels and the large channel system (*13*) is colored gray and was calculated using a probe radius of 1.4 Å. Numbered circles show points at which streamlines exit the protein and enter the lumen. These exit points are compared to the exits reported previously in Figure 4 of ref *13* which were denoted as "dashed arrow" exit, two "dotted arrow" exits, and "solid arrow" exit. Streamlines in the area of the back channel are colored orange with a purple exit point (point 1). Streamlines and exit points in the area of the broad channel are colored cyan (points 2 and 3). Streamlines and exit points in the area of the narrow channel are colored blue (point 4). Streamlines in the area of the large channel system and near the OEC are colored pink. Red spheres depict exits near the dashed arrow exit (points 5−8). Yellow spheres depict exits near the two dotted arrow exits (points 9−11). The green sphere depicts the exit near the solid arrow exit (point 12). The blue sphere depicts the exit from the large channel system leading into the narrow channel (point 13). The upper white sphere depicts the new exit near CP43-Thr412 (point 14). The lower white sphere depicts the new exit near PsbV-Asp79 (point 15). Insets A and B show snapshots at different times of the protein solvent-accessible surface and streamlines in two areas of PSII. Inset A shows the area near the two dotted arrow exits. Transient opening permits the access of water to the large channel system. Inset B shows the proximal end of the back channel. The streamline in the back channel continues all the way to the $CaMn_4$ cluster due to transient opening of the proximal end of the channel (see the text for details). The protein has been clipped close to this region to facilitate visualization of the opening. Residues previously proposed to act as a "control gate" in ref *13* are found in this region and are shown as licorice.

The large channel system discovered in ref *13* was reported to have four exits and identified as a preferred O_2 exit route. This suggested function was based on the idea that oxygen should be removed as soon as possible, and the large size and multiple channels and exits of this system make it well suited for this purpose. However, the possibility that water could also be supplied by the large channel system was not ruled out. The results of our simulation showed that the large channel system contains an abundance of mobile water molecules. All water streamlines found in the area of the channel system are shown in Figure 3D. Our results indicated multiple points of entrance for water into this system. How do these sites compare with the oxygen exit points previously identified from analysis of the static structure?

Group A includes streamlines in the area identified with a solid arrow exit in Figure 4 of ref *13*. Surprisingly, although this opening remained open and filled with water during the course of our MD simulations, we did not detect any streamlines propagating along this channel. We did, however, find streamlines entering the large channel system in this area near PsbV-Ser65 (point 12 in Figure 2), but as soon as streams passed this residue and entered the channel, the direction of water flow changed sharply. Behind the rather narrow opening of this channel, streamline tracing picked up several streams propagating in different directions, which would be expected in the wide and branched cavity in this region

Group B includes streamlines in the area with a dashed arrow exit identified in Figure 4 of ref *13*. We found a group of diverging streamlines with four water entrance points in this area. Two entrances are near PsbV-Tyr162 and -Tyr163 (points 7 and 8 in Figure 2) and two entrances near psbU-Lys134 (points 5 and 6 in Figure 2).

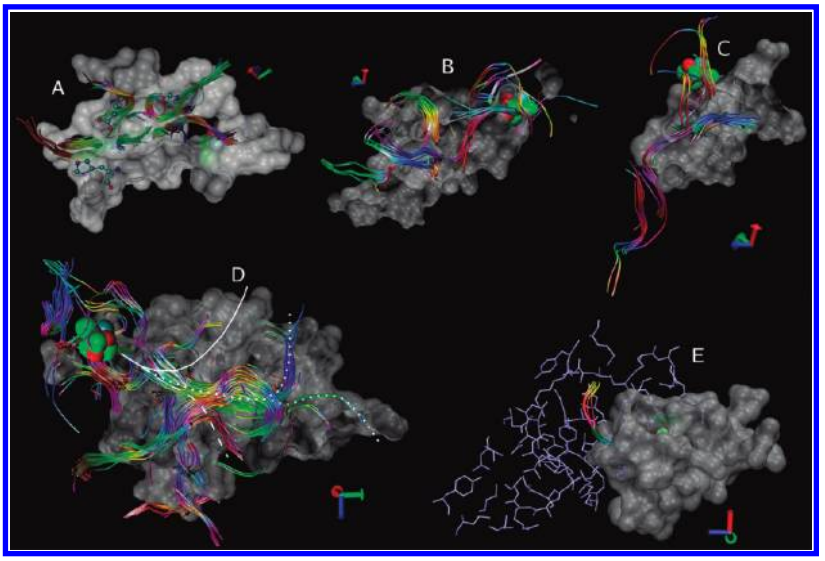

FIGURE 3: Streamlines in the areas of the broad, narrow, back, and large channels previously described in ref *13*. Gray surfaces show amino acid residues reported to line these channels and have been clipped at the level of the streamlines. The Mn, Ca, and O atoms of the OEC are depicted as green, blue, and red spheres, respectively. Streamlines are color coded by direction: red streamlines propagate along the *x* direction, green *y*, and blue *z*. (A) Back channel. The ball-and-stick representation shows two amino acids forming an exit from this channel (D1-His92 and CP43-Asp360) and one amino acid (CP43-Pro307) where the channel makes a sharp turn toward the OEC. (B) Broad channel. (C) Narrow channel. (D) Large channel. Solid, dashed, and dotted white lines show paths to the exits from the large channel system as described in ref *13*. (E) Water stream connecting the narrow (blue sticks) and broad (gray surface) channels.

Group C includes streamlines in the area of the two dotted arrow exits as identified in Figure 4 of ref *13*. We have found three water entrances in this region. Streamlines near PsbJ-Ser39 diverge and form two entrance points; the third entrance point is near D1-Ile307 and -Gly311, PsbJ-Leu40, and PsbV-Tyr52, -Lys56, and -Ile151 (points 9−11 in Figure 2).

Group D includes three streams leading to new water entrance points to the large channel system that were detected in our MD simulation. One of the streamlines (point 13 in Figure 2) enters the large channel system near PsbU-Tyr133 from the narrow channel. This entrance was transiently open to the surface. A second stream enters the channel system near CP43-Thr412 but does not connect to the surface of the protein (point 14 in Figure 2). The third stream enters the channel system near PsbV-Asp79 (point 15 in Figure 2). Two separate water streams lead to this point. One is a very short direct connection to the lumen; the other stream is longer, runs parallel to the membrane plane, and originates from an entrance point at the protein surface among CP43, PsbU, and PsbV.

Our MD simulations have shown that thermal motions of the proteins in PSII cause channel opening and closing. For example, Figure 2A shows the transient opening and closing of entrances (points 9−11) in the area of the two dotted arrow entrances reported in Figure 4 of ref *13*. All three of these entrances rapidly fluctuated between open and closed states; however, 90% of the time, at least one of them was open, permitting easy access of water to the large channel system. In addition, exits 2 and 3 in Figure 2 also switched rapidly between open and closed states, exit 2 remaining open 85% of the time and exit 3 remaining open 15% of the time.

To determine the relative use of the many potential water channels identified in our analysis, we followed the paths taken by all individual water molecules in our MD simulation that moved from the protein exterior to the OEC. During the 10 ns of our simulation, eight water molecules moved from the lumen to within 6 Å of the geometric center of the OEC. One water molecule took the broad channel via entrance 3. Seven water molecules used the large channel system, one each via entrances 7, 8, 10, 11, and 13, and two molecules used entrance 12 (Figure 2). It is clear that over the millisecond turnover time of water splitting, individual water molecules could easily find their way to the OEC via many different routes. In our simulation, the water moved into the protein through the "cracks" between polypeptide subunits and the supply of water to the OEC does not appear to be limited. However, controlling the activity of water or flow of water near the OEC to keep the $CaMn_4$ cluster stable may still be an important factor.

Identification of Potential Water Access Control Points. According to previous functional assignments for the channels, incoming water molecules would need to cross a gap between the back channel (assigned as a water channel) and narrow or broad channels (assigned as proton exit channels) to reach the active site of the $CaMn_4$ cluster. This gap (residues D1-Tyr161, -His190, -Asp189, -Phe186, and -Asn165 and the Ca^{2+} ion) was proposed to act as a control gate for regulating the access of substrate water to the $CaMn_4$ cluster through transient changes in the diameter of the opening (*13*). Many of the residues forming the gap have also been proposed to be important for water oxidation (*16*, *39*), indicating the possibility of differential water access to the $CaMn_4$ cluster during the S-state cycle. Our simulations did not detect water streamlines going from the back channel to the narrow or broad channels. However, we did find water streams inside the back channel extending beyond its previously found boundary toward the OEC to reach it directly (Figure 2B). This water stream is possible due to a transient widening of the proximal end of the back channel which opened this connection 20% of the time. The synergistic movement of the amino acids lining this part of the back channel results in an opening allowing water access to the active site. The proximal extension of the back channel is lined by D1-Gln165, -Glu189, -His190, -Asn296, -Asn298, -Asp342, and -Ala344 and CP43-Leu401, -Gly402, -Gly409, -Val410, and -Ala411; the locations of five of these are shown in Figure 2B. This region includes three of the five amino acids previously proposed to act as a gate between the

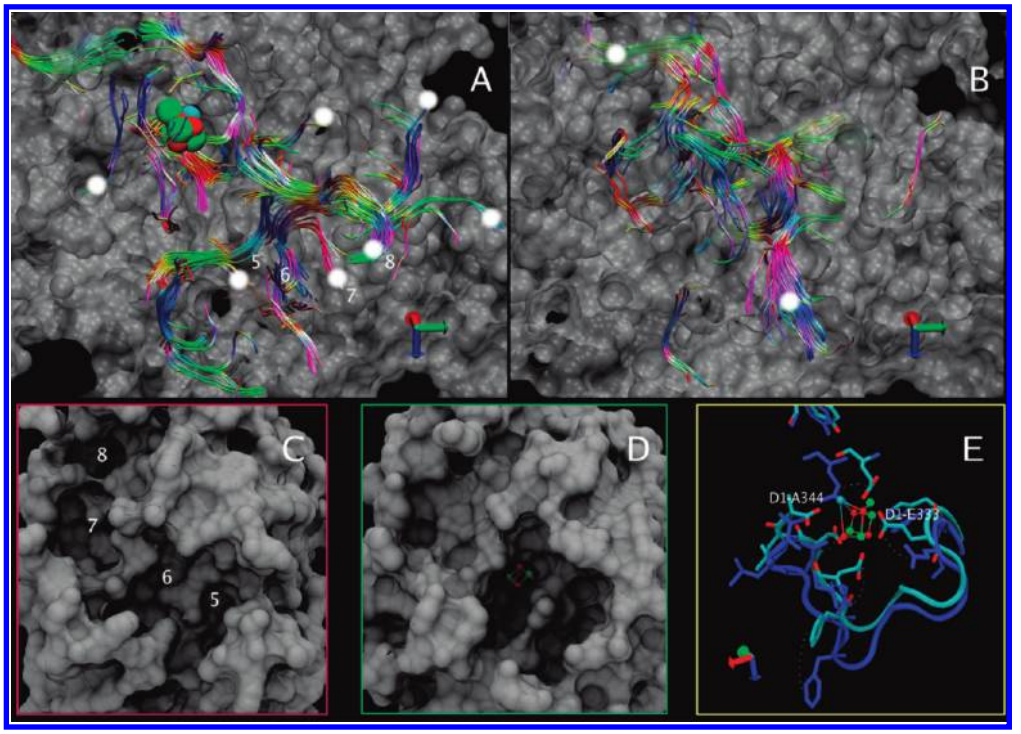

FIGURE 4: Comparison of the simulation with and without the OEC. Panels A (with the OEC) and B (without the OEC) show the streamlines, where the protein is colored gray and has been clipped below the streamlines for the sake of clarity. White circles show the most probable entrance points; in the simulations, at least one water molecule entered the protein via each of these points and ultimately reached the OEC sphere. In the simulation with the OEC, the water path is curved and branched with multiple water entrance points; without the OEC, the pathway is relatively straight with only two open entrances. Panels C (with the OEC) and D (without the OEC) show the surface of the protein in the region of the entrance point represented by the lower white circle in panel B. The view is along a line from that entrance point straight back to the OEC site. In panel C (with the OEC), although a number of channel openings are seen, the OEC is covered by protein and is not visible. In panel D (without the OEC), a "ghost" OEC has been placed at the OEC site to allow visualization of the straight and wide channel that goes from the lumenal surface all the way into the active site. Panel E shows the change in conformation of the protein in the region of the OEC site between the simulation in the presence of the OEC (gray licorice) and in the absence of the OEC (blue licorice).

back channel and narrow or broad channels. Our simulations confirmed these amino acids are important for water dynamics; however, as described above, the transient widening of the back channel revealed by our simulation is quite different from the previous proposal of a gate between channels (*13*). As a number of these residues have also been implicated in water oxidation, it would be interesting to study water flow in this region as a function of charge accumulation during the S-state cycle. The protonation states of titratable residues within the protein are expected to be affected by the changing charges on the $CaMn_4$ cluster. This may influence the dynamic protein structure and merits future investigation.

Water Streams in the Absence of the OEC. Although similar numbers of water molecules moved between the protein exterior and the OEC site in simulations with and without the OEC present, the pathways taken were different. In the absence of the OEC, the back channel and the large channel system were used exclusively and water entered from only two areas: point 1 and closely spaced points 5 and 6, which had fused into one entrance (Figures 2 and 4). The relatively straight water pathway in the simulation without the OEC is in stark contrast with the curved and branched water path in the simulation with the OEC (Figure 4). Many of the water entrance points observed in the presence of the OEC were transient and relatively small in contrast to the wider and continuously open entrances observed in the simulation in the absence of the OEC (Figure 4A,B). Differences between water streams with and without the OEC arose from changes in protein conformation in two regions: immediately adjacent to the OEC and an area where the C-terminal domains (CTD) of PsbU, PsbV, and D2 are in the proximity of each other and the surface of the protein. This was also the area of the highest probability for water to enter the large channel system in our simulation without the OEC [points 5 and 6 (Figures 2 and 4)]. In the absence of the OEC, we observed an outward displacement of the CTD of PsbU leading to the fusing of two narrow and transiently open entrances (points 5 and 6) into one wide and continuously open water entrance (Figure 4C,D). This process was initiated by a conformational change in the D1 CTD (Glu333−Ala344) which is stabilized by coordination of three amino acids, Glu333, Asp342, and Ala344, with the $CaMn_4$ cluster. When the cluster is absent, this loop shifts away from its initial position, dragging along the CTD of D2 (Figure 4E). This conformational change in the D2 CTD directly affects the CTDs of both PsbU and PsbV and initiates the opening of the gate. Interestingly, the amino acids near the CTD of PsbU, PsBV, and D2 are highly conserved between species.

DISCUSSION

X-ray structures of PSII revealed the OEC to be buried within the protein and thus begged the question of how substrate water might reach the active site and how oxygen and protons might leave. Previous work has searched for channels within the static structures, and continuous cavities reaching from the OEC to the lumen have been categorized and assigned functions on the basis of their relative size and hydrophobicity. However, without further information, such assignments remain arbitrary, and as previously stated by Ho et al. (*19*), dynamic studies are essential

for understanding channel function. Our molecular dynamic simulations of PSII have approached this problem by focusing on the location and movement of water. We found the lumenally exposed protein of PSII to be particularly spongelike and saturated, with more than 900 water molecules involved in a complex diffusion pattern. Water was found moving not only in previously identified water channels but also in so-called proton channels and regions that had never been identified as channels. During the 10 ns of our simulations, water molecules entered PSII at multiple sites and moved from the bulk water of the lumen to the OEC via different pathways.

Previous dynamic simulations of water conduction within proteins (aquaporin-1 and aquaglyceroporin) revealed that water molecules moved in a single file along a well-defined channel and that their translocations were correlated (*20*, *40*). In contrast, the water inside PSII was characterized by a complex hydrogen bond network within a highly branched channel system composed of wide and narrow areas and transiently connected hydrophilic pockets.

In our simulations, water effectively leaked into PSII from the lumen through multiple openings, many of which were transitory in nature and most of which were found in the cracks or interfaces between polypeptide subunits. This view of water movement within PSII is quite different from that generated by static cavity searches. However, it is clear that the redundancy offered by a manifold of interconnected water transport pathways through a "spongelike" region of lumenally exposed polypeptides might be the best way to ensure an adequate and robust supply of water to the OEC. By forming an interconnected network of nanopores, the protein also works effectively as a filter, keeping potentially harmful molecules away from the OEC and minimizing side reactions with small molecules carried with water. Incoming ions and cellular reductants are potentially harmful as they could impair water splitting by over-reducing the OEC or competing for binding sites.

In our simulation, water molecules moved between the OEC and the bulk water within 10 ns, much faster than the 1.6 ms turnover time for water splitting (*41*). This clearly indicates that delivery of water to the OEC is not limiting. The relatively fast delivery of water to the OEC has an impact on the interpretation of substrate oxygen isotope exchange measurements (*42*), eliminating the possibility that the slowest ^{18}O exchange rates were influenced by a protein barrier, as had previously been suggested in ref *43*.

Even though delivery of water to the OEC was not rate-limiting, we did observe restrictions of water movement in our simulation in the presence of the OEC. Restrictions were caused by transiently open pores both at the surface of the protein and in regions connecting different channel systems. This pattern of water flow is consistent with the idea of a kinetic barrier, required to slow movement of water near the $CaMn_4$ cluster to facilitate the ordered binding of substrate water and the formation of the O−O bond (*5*). Similar regulation of delivery of the substrate to the active site by the transient opening of passages in protein has been observed previously in acetylcholine esterase (*44*).

In stark contrast with the controlled access, transient openings, and multiple transport pathways observed in the presence of the OEC, our simulation of PSII in the absence of the OEC was characterized by formation of a single, relatively wider, straight path through PSII with two continuously open entrances. This change in channel geometry in the absence of the OEC is consistent with a requirement for conduction of Ca^{2+}, Mn^{2+}, and Cl^- ions necessary to rebuild the $CaMn_4$ cluster.

Our simulations show a significant difference in conformation in PSII in the presence and absence of the $CaMn_4$ cluster. Most of the difference was observed in the region of the C-terminal domains (CTD) of two extrinsic proteins, PsbV and PsbU. These regions were involved in formation of the wide opening at one end of the straight and wide channel in PSII lacking the OEC. Interestingly, comparison of PSII X-ray structures from *Thermosynechococcus elongatus* (PDB entries 1S5L and 2AXT or 3BZ1) shows a significant difference in the folding of the PsbU and PsbV CTDs. In the 1S5L structure, the two CT amino acids of both PsbU and PsbV form a flat and approximately round closed patch on the surface of the protein which would seal the access of water to the large channel system. Three of these amino acids are tyrosines which are well suited for gating (*44*). In the 2AXT and 3BZ1 structures, CTDs of PsbU and PsbV are not oriented to form a continuous surface patch. Our MD simulation in the presence of the OEC shifted the two CT tyrosines of PsbV to a conformation differing from either of the X-ray structures. These results suggest that a number of protein configurations in this region, characterized by different solvent accessibility, may exist at physiological temperature.

The potential involvement of PsbU and PsbV in regulation of ion transport in PSII is supported by a number of experimental observations. Both ΔPsbV and ΔPsbU deletion mutants show decreased oxygen evolution rates and loss of photoautotrophic growth under Ca^{2+} or Cl^- limiting conditions (*45*, *46*). In addition, the PsbV deletion mutant is prone to a rapid loss of oxygen evolving activity due to the loss of active site metal ions (*47*), and PsbU stabilizes oxygen evolution at elevated temperatures (*48*). All of these results are consistent with the active role of the CTDs of PsbU and PsbV in controlling water and ion transport observed in our PSII MD simulations.

CONCLUSIONS

Our results have demonstrated the benefits of streamline tracing for analysis and visualization of water flow in MD simulations. This technique is particularly useful for the identification of transiently formed channels in a complex system of nanopores. Our investigation of water movement in PSII has introduced a novel perspective to the study of the supply of water to the OEC. We have shown that functional PSII is characterized by a complex branched water supply structure with transient entrances and multiple control points. Loss of the OEC from PSII causes a transition to a simpler, straighter, and wider channel with two permanent openings that is well suited to the transport of ions required to rebuild the $CaMn_4$ cluster.

SUPPORTING INFORMATION AVAILABLE

Water streamlines in the areas of narrow, back, broad, and large channels in PDB format in orientation matching 2AXT X-ray structure. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Barber, J. (2003) Photosystem II: The engine of life. *Q. Rev. Biophys. 36*, 71–89.
2. Govindjee (2006) Photosystem II: The light-driven water: Plastoquinone oxidoreductase. Advances in Photosynthesis and Respiration (Wydrzynski, T. J., and Satoh, K., Eds.) Vol. *22*, Springer, Dordrecht, The Netherlands.

3. Joliot, P., Barbieri, G., and Chabaud, R. (1969) A new model of photochemical centers in system-2. *Photochem. Photobiol. 10*, 309.
4. Kok, B., Forbush, B., and McGloin, M. (1970) Cooperation of charges in photosynthetic O_2 evolution. 1. A linear four step mechanism. *Photochem. Photobiol. 11*, 457–475.
5. Wydrzynski, T., Hillier, W., and Messinger, J. (1996) On the functional significance of substrate accessibility in the photosynthetic water oxidation mechanism. *Physiol. Plant. 96*, 342–350.
6. Radmer, R., and Ollinger, O. (1983) Topography of the O-2-Evolving Site Determined with Water Analogs. *FEBS Lett. 152*, 39–43.
7. Force, D. A., Randall, D. W., Lorigan, G. A., Clemens, K. L., and Britt, R. D. (1998) ESEEM studies of alcohol binding to the manganese cluster of the oxygen evolving complex of Photosystem II. *J. Am. Chem. Soc. 120*, 13321–13333.
8. Anderson, J. M., and Chow, W. S. (2002) Structural and functional dynamics of plant photosystem II. *Philos. Trans. R. Soc. London, Ser. B 357*, 1421–1430.
9. Wraight, C. A. (2006) Chance and design: Proton transfer in water, channels and bioenergetic proteins. *Biochim. Biophys. Acta 1757*, 886–912.
10. Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J., and Iwata, S. (2004) Architecture of the photosynthetic oxygen-evolving center. *Science 303*, 1831–1838.
11. Loll, B., Kern, J., Saenger, W., Zouni, A., and Biesiadka, J. (2005) Towards complete cofactor arrangement in the 3.0 angstrom resolution structure of photosystem II. *Nature 438*, 1040–1044.
12. Murray, J. W., and Barber, J. (2007) Structural characteristics of channels and pathways in photosystem II including the identification of an oxygen channel. *J. Struct. Biol. 159*, 228–237.
13. Ho, F. M., and Styring, S. (2008) Access channels and methanol binding site to the $CaMn_4$ cluster in Photosystem II based on solvent accessibility simulations, with implications for substrate water access. *Biochim. Biophys. Acta 1777*, 140–153.
14. Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A., and Saenger, W. (2009) Cyanobacterial photosystem II at 2.9-angstrom resolution and the role of quinones, lipids, channels and chloride. *Nat. Struct. Mol. Biol. 16*, 334–342.
15. Gabdulkhakov, A., Guskov, A., Broser, M., Kern, J., Muh, F., Saenger, W., and Zouni, A. (2009) Probing the Accessibility of the Mn_4Ca Cluster in Photosystem II: Channels Calculation, Noble Gas Derivatization, and Cocrystallization with DMSO. *Structure 17*, 1223–1234.
16. Sproviero, E. M., McEvoy, J. P., Gascon, J. A., Brudvig, G. W., and Batista, V. S. (2008) Computational insights into the O_2-evolving complex of photosystem II. *Photosynth. Res. 97*, 91–114.
17. Sproviero, E. M., Gascon, J. A., McEvoy, J. P., Brudvig, G. W., and Batista, V. S. (2008) Computational studies of the O_2-evolving complex of photosystem II and biomimetic oxomanganese complexes. *Coord. Chem. Rev. 252*, 395–415.
18. Cohen, J., Kim, K., King, P., Seibert, M., and Schulten, K. (2005) Finding gas diffusion pathways in proteins: Application to O_2 and H_2 transport in Cpl [FeFe]-hydrogenase and the role of packing defects. *Structure 13*, 1321–1329.
19. Ho, F. M. (2008) Uncovering channels in photosystem II by computer modelling: Current progress, future prospects, and lessons from analogous systems. *Photosynth. Res. 98*, 503–522.
20. Zhu, F. Q., Tajkhorshid, E., and Schulten, K. (2001) Molecular dynamics study of aquaporin-1 water channel in a lipid bilayer. *FEBS Lett. 504*, 212–218.
21. de Groot, B. L., and Grubmuller, H. (2001) Water permeation across biological membranes: Mechanism and dynamics of aquaporin-1 and GlpF. *Science 294*, 2353–2357.
22. Chakrabarti, N., Roux, B., and Pomes, R. (2004) Structural determinants of proton blockage in aquaporins. *J. Mol. Biol. 343*, 493–510.
23. de Groot, B. L., Frigato, T., Helms, V., and Grubmuller, H. (2003) The mechanism of proton exclusion in the aquaporin-1 water channel. *J. Mol. Biol. 333*, 279–293.
24. Hu, Z., Jiang, J., and Sandler, S. I. (2008) Water in hydrated orthorhombic lysozyme crystal: Insight from atomistic simulations. *J. Chem. Phys. 129*, 075105.
25. Hu, Z. Q., and Jiang, J. W. (2008) Molecular dynamics simulations for water and ions in protein crystals. *Langmuir 24*, 4215–4223.
26. Vasil'ev, S., and Bruce, D. (2006) A protein dynamics study of photosystem II: The effects of protein conformation on reaction center function. *Biophys. J. 90*, 3062–3073.
27. Zhang, L., and Hermans, J. (1996) Hydrophilicity of cavities in proteins. *Proteins: Struct., Funct., Genet. 24*, 433–438.
28. Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., and Schulten, K. (1999) NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys. 151*, 283–312.
29. Wang, J. M., Cieplak, P., and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem. 21*, 1049–1074.
30. Simmerling, C., Strockbine, B., and Roitberg, A. E. (2002) All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc. 124*, 11258–11259.
31. Basma, M., Sundara, S., Calgan, D., Vernali, T., and Woods, R. J. (2001) Solvated ensemble averaging in the calculation of partial atomic charges. *J. Comput. Chem. 22*, 1125–1137.
32. Sproviero, E. M., Gascon, J. A., McEvoy, J. P., Brudvig, G. W., and Batista, V. S. (2007) Quantum mechanics/molecular mechanics structural models of the oxygen-evolving complex of photosystem II. *Curr. Opin. Struct. Biol. 17*, 173–180.
33. Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., and Aldroubi, A. (2000) In vivo fiber tractography using DT-MRI data. *Magn. Reson. Med. 44*, 625–632.
34. Zhukov, L., and Barr, A. (2002) in Proceedings of the 13th IEEE Visualization Conference (VIS'02), pp 387−394, IEEE Computer Society, Piscataway, NJ.
35. Lebihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., and Lavaljeantet, M. (1986) Mr Imaging of Intravoxel Incoherent Motions: Application to Diffusion and Perfusion in Neurologic Disorders. *Radiology 161*, 401–407.
36. Zaraiskaya, T., Kumbhare, D., and Noseworthy, M. D. (2006) Diffusion tensor imaging in evaluation of human skeletal muscle injury. *J. Magn. Reson. Imaging 24*, 402–408.
37. Basser, P. J., and Jones, D. K. (2002) Diffusion-tensor MRI: Theory, experimental design and data analysis—a technical review. *NMR Biomed. 15*, 456–467.
38. Johansen-Berg, H., and Rushworth, M. F. S. (2009) Using Diffusion Imaging to Study Human Connectional Anatomy. *Annu. Rev. Neurosci. 32*, 75–94.
39. Debus, R. J. (2005) in Photosystem II: The light-driven water: plastoquinone oxidoreductase (Wydrzynski, T., and Satoh, K., Eds.) pp 261−284, Springer, Dordrecht, The Netherlands.
40. Jensen, M. O., Park, S., Tajkhorshid, E., and Schulten, K. (2002) Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl. Acad. Sci. U.S.A. 99*, 6731–6736.
41. Dau, H., and Haumann, M. (2008) The manganese complex of photosystem II in its reaction cycle: Basic framework and possible realization at the atomic level. *Coord. Chem. Rev. 252*, 273–295.
42. Hillier, W., and Wydrzynski, T. (2004) Substrate water interactions within the Photosystem II oxygen evolving complex. *Phys. Chem. Chem. Phys. 6*, 4882–4889.
43. Hillier, W., and Wydrzynski, T. (2008) O-18-Water exchange in photosystem II: Substrate binding and intermediates of the water splitting cycle. *Coord. Chem. Rev. 252*, 306–317.
44. Zhou, H. X., Wlodek, S. T., and McCammon, J. A. (1998) Conformation gating as a mechanism for enzyme specificity. *Proc. Natl. Acad. Sci. U.S.A. 95*, 9280–9283.
45. Shen, J. R., Ikeuchi, M., and Inoue, Y. (1997) Analysis of the psbU gene encoding the 12-kDa extrinsic protein of photosystem II and studies on its role by deletion mutagenesis in *Synechocystis* sp. PCC 6803. *J. Biol. Chem. 272*, 17821–17826.
46. Shen, J. R., Vermaas, W., and Inoue, Y. (1995) The Role of Cytochrome C-550 As Studied Through Reverse Genetics and Mutant Characterization in *Synechocystis* Sp Pcc-6803. *J. Biol. Chem. 270*, 6901–6907.
47. Li, Z. L., Andrews, H., Eaton-Rye, J. J., and Burnap, R. L. (2004) In situ effects of mutations of the extrinsic cytochrome c_{550} of photosystem II in *Synechocystis* sp PCC6803. *Biochemistry 43*, 14161–14170.
48. Nishiyama, Y., Los, D. A., Hayashi, H., and Murata, N. (1997) Thermal protection of the oxygen-evolving machinery by PsbU, an extrinsic protein of photosystem II, in *Synechococcus* species PCC 7002. *Plant Physiol. 115*, 1473–1480.