Tracing the Pathways of Waters and Protons in Photosystem II and Cytochrome c Oxidase

Divya Kaur 1,2,†, Xiuhong Cai 1,3,†, Umesh Khaniya 1,3, Yingying Zhang 1,3, Junjun Mao 1, Manoj Mandal 1 and Marilyn R. Gunner 1,2,3,*

1 Department of Physics, City College of New York, New York, NY 10031, USA; dmatta@gradcenter.cuny.edu (D.K.); xcai@gradcenter.cuny.edu (X.C.); ukhaniya@gradcenter.cuny.edu (U.K.); yzhang5@gradcenter.cuny.edu (Y.Z.); jmao@ccny.cuny.edu (J.M.); mandalmanojcu@gmail.com (M.M.)
2 Ph.D. Program in Chemistry, The Graduate Center of the City University of New York, New York, NY 10016, USA
3 Ph.D. Program in Physics, The Graduate Center of the City University of New York, New York, NY 10016, USA
* Correspondence: mgunner@ccny.cuny.edu
† The first two authors contributed equally to this paper.

Received: 1 January 2019; Accepted: 28 January 2019; Published: 31 January 2019

Abstract: Photosystem II (PSII) uses water as the terminal electron donor, producing oxygen in the Mn₄CaO₅ oxygen evolving complex (OEC), while cytochrome c oxidase (CcO) reduces O₂ to water in its heme–Cu binuclear center (BNC). Each protein is oriented in the membrane to add to the proton gradient. The OEC, which releases protons, is located near the P-side (positive, at low-pH) of the membrane. In contrast, the BNC is in the middle of CcO, so the protons needed for O₂ reduction must be transferred from the N-side (negative, at high pH). In addition, CcO pumps protons from N-to P-side, coupled to the O₂ reduction chemistry, to store additional energy. Thus, proton transfers are directly coupled to the OEC and BNC redox chemistry, as well as needed for CcO proton pumping. The simulations that study the changes in proton affinity of the redox active sites and the surrounding protein at different states of the reaction cycle, as well as the changes in hydration that modulate proton transfer paths, are described.

Keywords: proton pumping; network analysis; proton affinity; proton coupled electron transfer

1. Introduction

The proton coupled electron transfer reactions that take place in the membranes of photosynthetic and aerobic organisms generate the transmembrane electrochemical gradient used to fuel ATP synthesis and other metabolic processes. Photosystem II (PSII) uses the energy of light and cytochrome c oxidase (CcO) the energy released transferring electrons from cytochrome c to O₂, to add to the proton gradient [1–3]. In PSII, the reduced quinone product retains enough energy to drive proton transfer catalyzed by complex III (the bc₁ or b₆f complexes), while CcO is the terminal redox acceptor, producing water as the end product of aerobic metabolism. Each of these proteins carries out proton coupled electron transfers, where the substrate changes protonation state as it undergoes redox chemistry. In addition, CcO also transfers protons across the membrane embedded protein from the N-side (negative side, high pH) to the P-side (positive side, low pH).

Water is a key constituent in PSII and CcO. Water is the substrate for PSII and product for CcO. Internal waters stabilize buried ionized groups and water pathways permit proton transfer between the surface and the protein interior. Thus, the reaction can be controlled by changing hydration levels through the reaction cycle. With the elucidation of the structures of these proteins, computational
simulations have made it possible to begin to understand the mechanism in atomic detail. Our focus will be on the simulation techniques that provide insight into the thermodynamics of electron and proton transfer reactions at buried sites in proteins and on the identification of essential pathways for protons and waters.

1.1. Photosystem II

Cyanobacteria, algae, and higher plants carry out oxygenic photosynthesis [1]. PSII uses water as the terminal electron donor, leading to the release of O$_2$ as a byproduct [4–6]. Within PSII, light activation leads to the oxidation of the redox active chlorophylls, P$_{680}^+$, to P$_{680}^+$ and electron transfer across the protein leads to reduction of plastoquinone (Q) to plastoquinol (QH$_2$). Oxidation of water takes place at the oxygen evolving complex (OEC), a Mn$_4$CaO$_5$ cluster. The redox active tyrosine, Y$_Z$, is the intermediate redox cofactor, reducing P$_{680}^+$, and in turn being reduced by the OEC [7]. The OEC is located near the lumen (P-side) of PSII so that vectorial electron transfer adds to the proton gradient across the thylakoid membrane without protons translocating across the membrane [2]. The overall reaction is

\[
\text{Lumen : } 2\text{H}_2\text{O} + 4h\nu \rightarrow 4\text{e}^- + \text{O}_2 + 4\text{H}_2^+.
\]

\[
\text{Stroma : } 2\text{Q} + 4\text{e}^- + 4\text{H}_2^+_\text{N} \rightarrow 2\text{QH}_2,
\]

Protons (H$_2^+$), are lost from water to the P-side while Q binds protons (H$_2^+_\text{N}$) from the N-side (stroma). Thus, Q reduction and water oxidation both add to the pH gradient. Overall, the transmembrane $\Delta p$H is increased by 1 proton/electron. Figure 1a shows a schematic picture of the PSII mechanism.

**Figure 1.** Schematic illustration of PSII and CeO showing key features of each mechanism. Red lines: electron transfer pathways; solid blue lines: chemical proton transfer pathways; dashed blue lines: pathways for pumped protons; black lines: substrate and product pathways. (a) In PSII the excited P$_{680}^+$, transfers an electron via a pheophytin (Pheo) and plastoquinone (Q$_\Lambda$) to a second plastoquinone (Q$_\beta$). Following a second turnover, the doubly reduced Q$_\beta$ has bound two protons from the N-side and dissociated into the membrane as QH$_2$. The active site oxygen evolving complex (OEC) comprised of four Mn, five $\mu$-O and one Ca in a cubane structure with a dangling Mn. It reduces P$_{680}^+$ via Y$_Z$. Following four turnovers, the OEC oxidizes H$_2$O to O$_2$ in the S state, releasing four protons to the P-side of the membrane. (b) In CeO, the active site, BNC, consists of heme $\alpha_3$ and Cu$_A$ with the redox active Y288. Electrons are transferred from cytochrome c via Cu$_A$ and heme a to the BNC. Following accumulation of four electrons, O$_2$ is bound to the BNC and reduced to H$_2$O. The four protons needed for chemistry are transferred from the N-side through the D- and K-channels. The four pumped protons enter through the D-channel and are transiently held in the proton loading site (PLS), located near the P-side of CeO, before being released to the P-side of the membrane. The D-channel ends at E286, which passes a proton to the BNC or PLS.
The OEC goes through five increasingly oxidized states (S0 to S4) losing four electrons [8,9] and generating a redox potential that is positive enough to oxidize water in the single step S4→S0 [10]. Much of the current research on the OEC focuses on producing a model of the S-state cycle that ends in O–O bond formation. Computational simulations [11,12] combined with XFEL structures [13,14] and many experiments [6] provide an emerging picture of the OEC structure through the S-state cycle and the mechanism by which the cluster can catalyze water oxidation. One question is how the OEC Mn cluster changes shape through the S-state cycle. Conventional crystal structures gave a glimpse [15], but the OEC becomes highly reduced (below S0) under X-ray illumination [16,17]. Crystal structures obtained with X-ray free electron laser (XFEL) measurements provide a chance to obtain information about individual S-state structures [14,18–21]. However, the protein in the crystal is never in one state: with conventional X-ray crystallography it is in a mixture of highly reduced states [17]; dark adapted XFEL structures will be mostly in the S1 state, but will contain some S0 as well; with multi-flash spectroscopy the non-unity quantum yield means there will always be a mixture of states [14]. In addition, extracting structural information for lighter-atoms (like O) surrounded by electron-rich Mn is problematic [22]; and the structures do not provide the location of protons [23,24].

The OEC is made up of three Mn, one Ca, and five μ-oxo ligands held together in a cubane structure with a dangling Mn (Figure 2a). The combinatorics of a cluster with four Mn lead to 4 possible redox isomers for S0 ((MnIII)2(MnIV)2), 4 for S2 ((MnIII)1(MnIV)3), and 1 for S3 ((MnIV)4). S4 removes an electron from a Mn or substrate oxygen [6,25]. Combined QM/MM analysis along with EXAFS and XRD measurements have established the redox isomer and atomic positions for S0 [26], S1 [27], S2 [28,29], and S3 [4,30–32] states. The low energy redox state for S0 is 3433 (giving the oxidation states of the Mn designated 1 through 4) while S1 is 3443 [33].

![Figure 2. Structure of the active sites of PSII and CcO. (a) The Mn4CaO5 cluster of the OEC. The water ligands to Mn4 and CaII as well as the nearby residues D1-E329, D1-H337, D1-D61, D1-E65, D2-K317 and nearby Cl are shown. (b) The structure of the BNC in CcO. CuB is the sphere. Heme a3, Y288 as well as the ligands of heme a3 and CuB are shown in sticks. Y288 is covalently bound to the CuB ligand H284, PRD and PRA are the propionate acids of heme a3 and may be part of the PLS. Substrate water will bind between CuB and the central iron of Heme a3.](image)

For S2, two redox isomers are found [29,33] as a result of Mn1 and Mn4 competing to make the bridging O5 their sixth ligand [29,34,35]. Model system studies [36] show MnIII can have five or six ligands but an oxidized MnIV needs six. In the redox isomer where Mn4 is oxidized (3444) O5 moves towards the dangler Mn4 giving a g = 2 multiline EPR signal. In the other, Mn1 is oxidized (4443) and O5 completes its coordination shell, giving the g = 4.1 broad EPR signal [37].
S₃ [38] is primed for reaction. The growing agreement about the structure of this state comes from a dialog between QM/MM simulation and XFEL experiment [20,22,30,39–41]. In the S₃ state, a new water (or hydroxyl) is added. The two S₂ isomers can lead to an open state for S₃ where the added substrate O is bound to the dangling Mn₄ or a closed state where the substrate is on Mn1. The incomplete ligand shells of Mn1 and Mn4 in states below S₃ may deny a position for the second substrate water to bind before S₃ [42–44].

There are four terminal waters in the OEC, two coordinated to Mn₄ (W₁, W₂) and two at CaII (W₃, W₄) (Figure 2a). It is proposed that substrate water may be contributed by a ligand to Mn₄ [4,41] or CaII [34,45], or it may arrive from a second shell water near O₄ [31,41].

1.2. Cytochrome c Oxidase

CcO (complex IV) is the terminal protein of the respiratory electron transfer chain in mitochondria and aerobic bacteria, using the electrons from cytochrome c produced in complex III, to reduce O₂ to water [46–49]. CcO catalyzes the reaction that reverses the OEC chemistry, using the energy liberated in the reduction of O₂ to add to the transmembrane proton gradient. However, rather than using a Mn based cofactor, the active site binuclear center (BNC) of CcO contains a heme (heme a₃) and a copper (CuB) [50] (Figure 2b). While the OEC is near the P-side of the protein, the BNC is in the center so that substrate protons must move into the protein, from the N-side, to produce water. During one catalytic cycle, four electrons are transferred from cytochromes c bound at the P-side, to Cuₐ, a bi-copper site, and then to a low-spin heme a, finally to the BNC [51]. Meanwhile, eight protons are taken up from the N-side, with four protons delivered to the BNC as needed for the oxygen chemistry (chemical protons), and four protons translocated across the membrane (pumped protons). The overall reaction is:

\[
O_2 + 4\text{Cytc}^{3+}_P + 8H_N^+ \rightarrow 2H_2O + 4\text{Cytc}^{3+}_N + 4H_P^+ ,
\]

The reaction leads to two charges/electron translocated across the membrane. Figure 1b provides a schematic illustration of the CcO mechanism.

CcO cycles through four increasingly reduced states (denoted O, E, R, and P/F) [48,49,52]. The P state occurs following reduction of O₂. P and F [53] have the same number of electrons, while F has bound another proton. As with the OEC where water oxidation is carried out in the S₁→S₂ step, CcO carries out the full reduction of O₂ in one step (R→P) to minimize the production of reactive oxygen species [51]. Heme a is the electron donor to the BNC, analogous to the relationship of Y₂ to the OEC. At each redox state transition, one electron and one chemical proton is sequentially added to the BNC to help ensure redox leveling. As in the OEC where the second water is not added until late in the reaction cycle (in S₃), O₂ is only bound when the BNC is in the fully reduced, R, state (Fe₃⁺, CuB, Y–OH). The ensuing reaction then oxidizes the BNC and reduces O₂ to form the P state (Fe⁴⁺V, CuB⁻, Y–OH⁻, YO⁻). Thus, the bound O₂ accepts four electrons, two from the feryl heme a₃, one from CuB, and the fourth from a tyrosine covalently bonded to one of the three histidine ligands of CuB. This yields the Pₐ state [54]. O₂ reduction can alternatively pass through the Pₐ state (Fe⁴⁺V, CuB⁻, Y–OH⁻, Y–O⁻) when the reduced heme a rather than the tyrosine contributes an electron to the reaction [55]. The subsequent reduction of the BNC is coupled to the addition of protons to the oxygen bound to heme a₃ and CuB, which are released as water [53,56,57]. The currently accepted order of proton and electron additions is F: (Fe⁴⁺V, CuB⁻, Y–O⁻) to O: (Fe⁴⁺V, CuB⁻, Y–O⁻), to E: (Fe⁴⁺V, CuB⁻, Y–OH⁻, CuB⁻, Y–O⁻), and to R: (Fe⁴⁺V, CuB⁻, Y–OH⁻, CuB⁻, Y–O⁻). The catalytic cycle of CcO has been established by DFT simulation combined with many experiments [51,53,58–60].

2. Finding the Protons in the Active Site and the Surrounding Protein

PSII and CcO both carry out multielectron chemistry, where a series of single electron transfers create a sequence of redox intermediate states. The loss or gain of electrons is coupled to the loss or gain of protons [61]. The order of oxidation of the four OEC Mn or the BNC heme a₃ and CuB reflects
both the relative redox potential of each site as well as the proton affinity of the different protonatable groups in the cluster. Protonation changes of products or substrates (chemical protons) are driven by large changes in proton affinity due to oxidation or reduction of the Oxo–Mn [62–64] or heme–Cu redox cofactors [65].

In addition, there are many buried acidic and basic amino acids in the protein surrounding the various redox active sites. The charge state of these residues can modulate the electrostatic potential, tuning the OEC and BNC electron and proton affinity. These residues may change protonation states in response to changes at the active site. The proton transfer pathways to buried sites also require buried groups that can transiently change protonation state. The protonation and tautomer state of both the active site and the protein is often experimentally elusive, yet the thermodynamic analysis of the redox cycle relies on establishing the correct states for many sites in the protein and cofactor. In addition, any molecular dynamics (MD) or QM/MM simulation must make a best guess as to which groups are protonated to start the calculation. There are several simulation techniques that can help establish the likely protonation or tautomer states by the analysis of the proton affinity of the cofactor and protein in each redox intermediate state [66–69].

2.1. Methods of Calculation

An acidic or basic residue can be in its protonated or deprotonated form and each redox cofactor has a defined oxidation state. A microstate of the protein will define the protonation state of each residue and the redox state of each cofactor. If each of the n groups of interest has m protonation or redox states, there are \( m^n \) microstates. In the calculation of Boltzmann distribution of protonation states within the protein, it is generally assumed that each intermediate in the redox cycle finds the lowest energy protonation states and redox state isomers and that the protonation state of the surrounding protein can come to equilibrium with the redox changes within microseconds. There are a variety of methods to determine the relative free energy of the possible redox and protonation microstates [66,67]. The different simulation methods differ in the level of theory used, the elements that are fixed or allowed to change and the region of the protein that is included in the simulation. The key methods that help to establish the redox and protonation equilibria distributions are QM based calculations (DFT and QM/MM), CE/MC methods that use Monte Carlo (MC) sampling with continuum electrostatics (CE) energies and constant pH molecular dynamics (cpHMD). These methods are distinct from conventional molecular dynamics (MD) calculations, as these must fix the chemical form of the protein and cofactors.

Monte Carlo approaches that use classical CE energies allow the sampling of multiple redox states and many protonatable groups in the whole protein [66]. The programs MCCE [70] and Karlsberg [71] use this approach. These methods do not calculate \( E_m \) or \( pK_a \) from first principles. Rather a known titration of an analog of the residue in water provides a reference energy [64,65,72]. The calculations then determine how interactions within the protein modulate the energy, shifting the electron and proton affinity. Assuming we have a good value for the interaction energy between all pairs of charges, the equilibrium distribution of protonatable residues and redox active cofactors is determined via Monte Carlo sampling. The resulting Boltzmann distribution at a given pH and \( E_{th} \) can have groups with a non-integer probability of being protonated or reduced as would be found in experiment near the \( pK_a \) or \( E_m \). The use of Grand Canonical Monte Carlo (GCMC) methods allows the binding or release of functionally important ions, such as chloride, and of explicit waters in the same analysis that determines the distribution of ionization states.

One strength of CE/MC methods is that they can reliably converge to the equilibrium distribution for the whole protein. It can be relatively straightforward to interrogate the outcome to determine the contribution of individual residues to the result. A strength and weakness of the CE method is the use of a dielectric constant. This accounts for the different polarization of the protein, water, and lipid that screens interactions, so provides an inexpensive polarizable force field. However, the value of the dielectric constant is an ongoing subject of dispute [66]. CE methods can only include the influence of
quantum effects such as changing spin state or of Jahn–Teller effects by having access to values for $E_{m}$s and $pK_a$s of closely related reference reactions. It is also difficult to incorporate protein motion into the calculation because MC sampling is far more efficient with a fixed protein backbone. The calculations may allow side chain rotamer sampling [69,70]. A current work-around to the fixed backbone is to average the charge state distributions calculated in multiple snapshots from MD simulation or from multiple crystal structures, which provide a distribution of likely conformations. One problem is that an MD simulation will always seek to stabilize the charge states assigned to the simulation [73,74].

The goal is to analyze the free energy of different mixtures of protonation and redox states while the protein samples conformational space. Constant pH MD (cpHMD) is an emerging technique which may be able to accomplish this goal [75–80]. One method is to run MD for a short time and then exit and use MC to determine the equilibrium protonation state and then resume the MD trajectory. While this method has significant promise, it has serious barriers to its application to the large proteins of interest here. Each addition or removal of a charge from the protein in a MC step creates a major disturbance for the system when resuming MD, requiring re-equilibration after each move. This process is quite difficult to converge, especially for the proteins of interest here where many sites have closely coupled changes in ionization states. In addition, the classical MD portion of the simulation uses Coulomb’s law in vacuum, which can overestimate the electrostatic interactions [81]. As the key questions involve the interactions of buried charges the uncertainty of the electrostatic energy may be addressed by use of polarizable force fields. Unfortunately, this adds significant computational cost [80].

Consideration of the OEC and BNC requires analysis of multi-electron processes in biometallic complexes [82]. Simple classical methods cannot account for the changes in electron distribution, spin states, polarization, and cluster geometry that accompany the reactions. These elements require a quantum mechanical (QM) approach. To permit the calculations to include sufficient number of atoms DFT [83] and other semi-empirical [84] methods are used. The accuracy and computational cost of all QM calculations also depend on the basis set used. QM/MM (quantum mechanical and molecular mechanical) approaches allow the effects of the surrounding protein to be included [85]. All of these QM based methods calculate the optimized geometry and total energy for the system with a fixed number of protons and electrons [6,11,12,86,87].

2.2. Changing Protonation States of the OEC through the S-State Cycle

The mechanism of $O_2$ evolution involves both electron and proton transfer reactions [88–90]. Over the course of the reaction, cycle protons must be shed from the substrate waters. Coupled electron and proton transfers also keep the OEC catalyst from building up excess charge. The redox potential of a Mn with a protonated $\mu$-oxo bridge or neutral terminal water ligand will be more positive than when the proton is lost [62–64]. The oxidation of $\text{Mn}^{\text{III}}$ to $\text{Mn}^{\text{IV}}$ lowers the $pK_a$ of a terminal water ligand by as much as ten pH units [62–64]. Thus, the proton distribution will be strongly influenced by which Mn in the OEC is $\text{Mn}^{\text{III}}$ or $\text{Mn}^{\text{IV}}$ and conversely the oxidation state of the Mn will be influenced by the protonation state of the bridging oxygens and bound waters.

Experiments show one proton is lost between $S_0$ and $S_1$ and between $S_2$ and $S_3$ while 0 [91–94] to 0.3 [24] are lost on formation of $S_2$. There is general agreement that in $S_0$ there is a proton on the oxygen bridging Mn1 and Mn4 (O5) that is lost forming $S_1$. This leaves the terminal waters as having the only protons bound to the OEC beyond $S_0$. There is also general agreement that W3 and W4 on the $\text{Ca}^{\text{II}}$ will remain neutral in $S_0$, $S_1$, and $S_2$ [5].

The protonation states of W1 and W2 bound to Mn4 remain an open question. The Mn–O bond length in XFEL structures supports these waters being neutral in the well resolved $S_1$ (3443) structures where Mn4 is in the $\text{Mn}^{\text{III}}$ oxidation state [21]. FTIR spectroscopy suggests formation of a Mn-bound hydroxyl in the multiline, 3444, $S_2$ state [95]. FTIR spectroscopy [96] and XFEL structures [97] support W1 and W2 being neutral in the 4443 $S_2$ isomer, while there being one hydroxyl and one water in the 3444 $S_2$ isomer [98].
DFT and QM/MM calculations must define the number of protons in the system and a number of different choices have been made. Calculations have used a hydroxyl on W2 of Mn4 in S1 [99] and S2 [32,100], fully protonated water in S1 [101] and hydroxyl on W2 in S2 [23] or fully protonated water in both S1 and S2 states [28]. Recent classical MCCE calculations, using a CE/MC method, found that at equilibrium there is a mixture of OECs with W2 being hydroxyl or water in the multi-line, 3444 S2 state [99]. We are not referring to this situation as W2 being at its pKₐ as the pH dependence does not follow the Henderson-Hasselbalch equation. The finding that Mn4 bound H₂O and OH⁻ are close in energy may help explain the difficulty in defining the protonation state of W2 in this redox state. Thus, oxidation from S1 to S2 moves the proton affinity of W2 so that protonation hangs in the balance. There is little penalty to hold on to the proton, leading to little proton loss on forming S2. However, there is little energy needed to lose a proton, which is needed as the OEC moves into the S3 state [21,31,41].

2.3. Changing Protonation States of Residues in the Protein through the Reaction Cycle

Changes in the protonation state of the OEC substrate water or the water produced in the BNC are required to complete the reaction. In addition, each reaction cycle induces changes in the protonation and conformation states of the rest of the protein [6,51,99,103]. Some of these changes are essential to the function as they lead to proton transfer away from the N-side and into the P-side of the protein in the membrane. CE/MC methods, which allow the whole protein to come to equilibrium with each protein redox state intermediate, have been used to evaluate the changes in protonation states through the S-state cycle in PSII [104] and through the reaction cycle in CcO [53,105]. It should be noted that introduction of a charge will shift the proton affinity of all nearby residues. However, to release a proton the pKₐ start above the pH and shift to be below it; conversely proton binding requires the pKₐ to shift from below to above the pH [3].

2.3.1. PSII Protonation State

The protonation of all residues in PSII were determined with the OEC in different S-states using the CE/MC program, MCCE [102]. The S-state structures optimized by QM/MM [11] were docked into the 3ARC structure [106]. Polar hydrogen positions were sampled. Explicit waters were included and subjected to GCMC sampling. Of the acidic and basic residues within 15 Å of the OEC, only D1-E329 and D1-H337 change protonation during the transition from S0 to S2. There is a 50–80% probability that W2 will lose a proton going from the S1 to S2 state [102]. D1-E329, in the large channel described below, has a pKₐ near the pH, so it increases its probability of having a proton bound as W2 is deprotonated. In addition, there is a small probability of proton transfer from W1 to D1-D61. The stability of the protonation states of groups near the OEC is highlighted when the pH is raised from 6 to 8 in the MCCE analysis [102]. Calculations show no protonation state change in any residue within 15 Å of the OEC except for D1-E329. At pH 6, D1-E329 acts as buffer during the S-state transition while it is fully deprotonated in all S-states at pH 8.

2.3.2. CcO Protonation State

In CcO protons need to be translocated across the membrane and thus residues must experience ionization state changes coupled to the BNC redox and protonation state cycle. There is a proton loading site (PLS) that transiently binds protons for release, which has been assumed to lie on the P-side of the protein relative to the BNC. The proton affinity of the PLS would increase when the protein is in the BNC states associated with proton loading and decrease to unload protons as the cycle progresses. A number of residues had been suggested to be the PLS including the propionate A [51,107] or propionate D [108] of heme a₃, or H334, which is one of the ligands of CuB [109,110].

The stochastic MC analysis in MCCE was used to evaluate the identity of the PLS in an unbiased way [111]. Each of the four additions of an electron and a proton to the BNC through the P/F, O, E, R states is coupled to one proton pumped. Thus, the MCCE calculations broke the electron and proton transfers involving Heme a, E286 and the BNC that occur in each redox state transition into
six steps and the protonation states of all residues in the protein were determined. Observed changes
in the probability of protonation of several heme propionic acids led to the identification of the PLS
as a cluster of residues, instead of a single site. The detailed division of the reaction cycle shows the
proton is loaded to the PLS when heme $a$ is reduced, while the PLS releases its proton when a proton is
transferred from E86 to the BNC. These results amplified previous models of CcO function [51].

3. Finding Water and Proton Transfer Pathways

Protons must be transferred through protein in a step-wise manner, with a proton donor passing
it to a hydrogen bonded proton acceptor. Water, with a lone pair to accept a proton and a bound
proton to donate, plays an important role in proton transfer via the Grothuss mechanism. Thus,
potential proton transfer pathways are identified as consisting of hydrogen bonded chains of water,
with protonatable residues providing sites for metastable proton transfer intermediates. The rate of
proton transfer will depend on the probability that a pathway is connected and on the free energy of
injecting a proton into the path.

3.1. Methods to Find Water and Proton Pathways

There are several methods to find proton pathways. The simplest is to find connected cavities
with programs such as Caver [112] or Hole [113]. Other programs such as Dowser++ carry out a
semi-empirical analysis to evaluate the likelihood that waters will reside in these cavities [114]. These
programs analyze a single protein structure and determine if there is sufficient room for water, with a
defined radius (generally 1.4 Å). The simplicity of these methods makes them ideal for the first analysis
of a crystal structure.

MD trajectories for protein embedded in explicit water will show many waters traveling within the
protein. Grand Canonical Monte Carlo (GCMC) sampling also allows waters to come to equilibrium,
moving between the protein and solvent, generally in a structure with a fixed protein backbone. Both
MD and GCMC methods use classical force fields to determine the probability of water being at
each position.

The waters in MD trajectories can be analyzed in numerous ways. For example, the distribution
of the number of waters around a key residue through the trajectory or the probability that proton
donor and acceptors sites will be connected via waters can be found [115]. Potential of mean force
(PMF) methods allow us to determine the energy of waters or protons passing through a specified
channel [116,117]. The calculations start from a MD trajectory and determine the energy of the particle
at each location along the path. This can find the barrier for transfer as well as the low energy positions
where a water or proton may be trapped. One disadvantage is that the pathway to be evaluated is
constrained to path(s) found in the initial MD trajectory. Also, as this uses classical MD methods, the
protonation states are fixed and cannot come to equilibrium with any changes as the particle moves
along the path.

Markov state models (MSM) [118] have been applied to analyze which dynamic motions are
coupled to water cavity opening. Here at each point in a trajectory the cavity hydration is correlated
with a few other collective geometric variables [118]. The constructed MSM trajectories are projected
into the collective variable space, and clustered into microstates, then into several macrostates, which
vary from a dry and compact cavity macrostate to one that is wet and expanded. The MSM transition
matrix can be used to compute the mean first passage times (MFPTs) of the opening (to wet and
expanded) and closing (to dry and compact) cavity transitions [118].

Hydrogen bond connections made in many snapshots in an MD trajectory or the many accepted
microstates of an GCMC analysis can be subjected to a network analysis [119]. Programs such as
Cytoscape [120] and Gephi [121], permit visualization of the network (see Figures 3b and 4). MCCE [70]
carries out GCMC analysis of the water orientation and occupancy and the protonation state and
proton positions of residues in the whole protein. Network analysis [119] then generates a hydrogen
bond matrix for the resultant Boltzmann distribution. Long-range connections between residues
bridged by water chains are found. The hydrogen bond network is gathered into highly connected clusters. The model is that protons can move within these well-connected buried clusters, populating the sites with the highest proton affinity. The importance of different levels of hydration is seen when inter-cluster connections are found to change in MD or crystal structures prepared in different redox or protonation states. The GCMC sampling of the hydrogen bond network is an efficient method to determine the many connections that are available even with a single protein backbone, while considering possible changes in protonation states. However, as proton transfers through proteins occur relatively slowly, the formation of working proton pathways may be rare events which would often be missed by their being disallowed in the input structure.

3.2. Water and Proton Transfer Pathways in PSII

Water channels regulate the movement of protons, water and O$_2$ between the OEC and the P-side lumen. The current picture is that three channels play vital roles in this transport [122] (see Figure 3a,b), although other channels have been seen. The “narrow channel” [123] is formed by PsbO and PsbU subunits connecting D1-D61 and the µ-O4 ligand of the OEC to the lumen while the “broad channel” [123] is on the other side of D1-D61 exiting at the interface of D2 and the PsbO subunits [122]. The “large channel” has multiple branches exiting at the interface of the PsbV/PsbU subunits. The narrow channel is likely to conduct water [11,124] while the broad channel is the likely proton exit path [122,125–127] and the large channel may deliver substrate water [122]. For simplicity, we will use this nomenclature for the channels despite their being given a variety of names in earlier publications [44,123,128,129].

![Figure 3. (a) Water channels around the OEC. The QM/MM hydrated, optimized S$_1$ cluster is docked into the 3ARC [106] PSII structure. Red: large channel, pink: narrow channel; green: broad channel. (b) Hydrogen bond connections for residues in narrow, broad and large channels that are connected through cluster of water molecules as determined by MCCE network analysis [119].](image)

Initial methods found cavities by geometric analysis of the early, low resolution 1SSL [15] and 2AXT [130] structures [123,128,131]. The cavity discovery programs found the narrow and broad channels. In addition, they identified another hydrophilic (large) channel and a hydrophobic (back) channel proposed to carry O$_2$ [123,131].

The functionality of the geometry based studies were tested by PMF multiple steered molecular dynamics (MSMD) simulations [132]. Water molecules were continuously injected near Ca$^{II}$ and Mn$^{4}$, the likely binding sites for substrate water molecules. The calculated energetics of water movement indicate only the narrow, broad, and large channels have low barriers for water permeation. The proposed back channel [123,131] is permeable to O$_2$ but is inaccessible to water.
Continuum electrostatics/Monte Carlo (CE/MC) studies \[126\] were carried out on the 2AXT structure \[130\] to calculate the $p$ of D1-D61, D1-E312, D2-K317, PsbO-D158, PsbO-D222, PsbO-D224, PsbO-H228, and PsbO-E229 in the broad channel, using the Karlsberg program \[71\]. The $pK_a$s were found to increase monotonically moving from the OEC to the lumen. This observation led to the proposal that the broad channel will conduct protons outward, with the gradient of the proton affinity slowing proton leakage back into the protein.

Recently, combined MD and continuum electrostatics studies on 3ARC \[106\] cyanobacterial PSII and 3JCU \[133\] of spinach PSII showed how the properties of the large channel in cyanobacterial PSII (running through subunits denoted PsbU/V) and spinach PSII (in PsbP/PsbQ) are conserved \[134,135\].

The broad channel also runs through the PsbO subunit in PSII. In the broad channel, pairs of acidic amino acids surrounded by water clusters have been proposed to facilitate the proton transfer to the lumen \[136\]. A truncated PsbO-β subunit was crystallized at pH 6 and pH 10. MD simulations were carried out on the two structures with D102, which is in a pair with E97 in the loop of PsbO barrel of PSII, in different protonation states \[125\]. These two acidic residues are conserved. D102 is fixed in its protonated form in the pH 6 MD simulation and deprotonated in the simulation run on the pH 10 structure. All other carboxylates are ionized in both MD simulations. When D102 is protonated D102-E97 associates with one and five waters in an orderly manner that can conduct protons via a Grotthuss-type mechanism. Thus, the cluster may trap a proton preventing its backflow to the OEC. The cluster is disrupted at pH 10 when both carboxylic acids are ionized.

D1-D61 and D2-K317 are both found at the entry to the broad channel, near a Cl found in the 3ARC PSII structure \[106\]. Experiments \[137\] show that Cl is required for the OEC to progress beyond the $S_2$ state. MD and MCCE simulations \[138\] show that Cl removal results in the formation of a salt bridge between D1-D61 and D2-K317 blocking proton transfer through the broad channel \[139\]. In addition, MCCE calculations suggest the loss of the Cl \[138,139\] will change the protonation states of residues near the OEC and increase the number of protons lost when the OEC is oxidized from $S_1$ to $S_2$ \[102\].

There are several physical methods that can help follow the pathway of water, $O_2$, and protons \[123,128,129,132\]. Radiolytic footprinting can trace the residues surrounding water channels and also buried waters \[140,141\]. Exposure of the protein to X-rays produces OH· which will oxidatively modify nearby amino acids. These are then identified by mass spectrometry. Modified residues are found along the broad and large channels \[140,142\].

Mutagenesis studies have also modified the residues around the OEC to understand their role in assisting proton exit during the S-state transitions. D1-D61, which is hydrogen bonded to the W1 ligand of Mn4 is at the entry of the narrow and broad channels (Figure 3b). It can accept protons released from the OEC during water splitting and pass them into these channels. D1-D61N \[143\] lowers the $O_2$ evolution efficiency. FTIR spectroscopy on D1-D61A \[144\] shows changes in the hydrogen bond network around the OEC. MCCE studies \[102\] on D1-D61A mutants show more proton loss to the solvent on the $S_1$ to $S_2$ state transition relative to that found with the wild-type protein.

Another residue that is also part of the proton exit pathway is D2-K317. FTIR spectroscopy \[145\] on the mutants D2-K317A, K317R, K317Q, and K317E find the hydrogen bond network is perturbed. The FTIR measurements suggest conformational changes in the protein backbone that may help disrupt proton movement.

3.3. Water and Proton Transfer Pathways in CcO

CcO is a proton pump, translocating protons from N- to P-side of the membrane through water filled channels and cavities. It requires gates on both N- and P-side to prevent protons from traveling in the wrong direction. The interior of CcO on the N-side is relatively non-polar, so that the waters and polar residues in the two proton channels (D and K) are easy to see in the crystal structures. The functionality of residues in these channels has been supported by the impact of their mutation on the function \[146\]. The D-channel transports both chemical protons to the BNC and pumped protons to the PLS, while the K-channel carries only chemical protons \[147–149\] (see Figure 1b). While the
N-side channels are visible in crystal structures, no path from the D-channel to the PLS is found in the crystal structures. As will be described below MD simulations varying the protonation states of key residues are beginning to show the proposed connections.

3.3.1. The P-Side Exit Path

In contrast to the N-side, the P-side of CcO contains many buried polar residues that surround the PLS. Thus, it was not possible to identify the pathway to the P-side by inspection. Electrostatic calculations on bovine CcO, suggested that K171/D173 in subunit II to be promising candidates to lie on the proton exit pathway. These well-conserved residues had neutral and ionized forms close in energy as determined by CE/MC analysis so could bind and release protons with little change in energy [150]. Inspection of the Bovine CcO structure suggested an H-channel traveling from N-side to P-side [49,151,152], supported by mutagenesis [152]. However, mutations of the analogous H-channel residues in *Rb. sphaeroides* CcO [153] and yeast [154] did not support this assignment. Recent simulations with MCCE hydrogen bond network analysis in *Rb. sphaeroides* CcO [119] proposed a different proton exit pathway via N96, T100, S168 (*Rb. sphaeroides* numbering) with the exit close to H93/E182.

3.3.2. The Path through the Protein

In CcO, channels must be able to open and close to ensure that protons are bound from the N-side and released to the P-side. The hydration level of a cavity can serve as a switch, where the hydrated cavity promotes proton translocation and the dehydrated cavity blocks the proton path [51]. One of the key features of CcO is an orderly decision to transfer protons to the PLS for pumping or to the BNC for chemistry, both via the D-channel. Crystal structures, while containing many waters, often appear to be relatively dry. Thus, the CcO crystal structures do not show connections from the D-channel to the PLS, ~10 Å away, or to the BNC.

MD analysis found that protonating one of the heme propionic acids causes the rearrangement of a loop, which is anchored by a hydrogen bond to the ionized propionic acid. The loop opening creates a cavity (the Glu cavity) that fills with water in nanoseconds in a MD simulation [108,118]. This cavity stabilizes the charge on the isolated E286 (see Figure 1b), which is a likely proton transfer site [108]. E286 has a pKₐ in the dry, resting protein over 10 [105]. The coupling of the conformational changes of key residues to cavity opening were characterized by MSM analysis, allowing key hydrogen bond connections to be identified [118].

MD snapshots of *Rb. sphaeroides* CcO were subjected to MCCE network analysis [119]. Well hydrated MD snapshots show connections from a polar residue cluster on the N-side of the protein through the D-channel to the PLS and on to the P-side of the protein (Figure 4). However, the K-channel remains disconnected [155]. The exit of the hydrogen bond network [119] is near several identified water exit pathways [118]. The network analysis was also carried out on snapshots from MD trajectories run changing the protonation states of key residues that lead to changes in hydration. The highly interconnected clusters (shown as circles in Figure 4) are always present, while the connections between clusters are made and broken by changes in hydration. The Glu cavity must be hydrated to form connections between the D-channel and both the BNC and PLS [119]. A small water cavity, which may be controlled by the hydrophobic pair of I99 and L255B, is proposed to be able to open and close the proton path between the PLS and the P-side preventing back flow [119].
Figure 4. Hydrogen bond network in *Rb. sphaeroides* CcO [119]. The network is divided into circular clusters with well-connected residues shown as: N-surface, D-channel, K-channel, BNC, PLS, PLS’, P-exit. Residues are shown as individual nodes, labeled by residue number, followed by B to identify residues in subunit II. Heme a3 is labeled as HE3, while CuB as CUB, propionic acids of heme a3 as PA3/PD3, propionic acids of heme a as PA/PD. The labels are colored based on amino acid type; red: acid; blue: base; green: neutral polar; orange: propionic acid; purple: BNC. The larger nodes and darker edges show more important connections in the network analysis.

There has been a discussion about the role of the Glu cavity hydration for controlling proton transfer from the D-channel to the PLS. One critical question is how such a well hydrated cavity can block leaks from the P-side back to the D-channel or the BNC [48]. In addition, while the propionate D of heme a3, triggers the hydration of the cavity, propionate A of heme a3 may have a higher proton affinity. The question arises if this proton transfer to propionate A would rapidly close the cavity [156]. However, in the MD simulations while the cavity opens quickly, it closes more slowly so the shift of a proton within the PLS may lead to only slow dehydration [118].

The orientation of water chains within the channels has also been proposed to act as a switch to control proton transport, promoting unidirectional proton pumping. It has been suggested that the electric field between heme a and the BNC will orient water, favoring Grothuss proton transfer in one direction, especially when there is a single line of waters in a relatively dry channel [157]. However, the electric field is unlikely to be strong enough to explain the directionality of the proton pump [48,86] and give the slow rate of water reorientation after transfer of one proton the overall transfer rate may not support efficient CcO turnover by this mechanism [48,158].

3.3.3. Testing the Location of Water Channel Experimentally

Hydrogen–deuterium exchange measurements have been used to monitor the residues that have greater access to the solvent during turnover in CcO [159]. Various regions of the protein backbone are identified that may run along proton channels. Regions in the proposed hydrogen bond network do show enhanced exchange in a functioning CcO [119]. One interesting observation is that under conditions when the K-pathway is assumed to be active the D-pathway shows reduced H/D exchange, suggesting a mechanism that can provide alternating access to the two channels.

4. Conclusions

The OEC of PSII catalyzes the difficult oxidation of water using only earth abundant elements so that water, the universal biological solvent, can be the terminal electron donor. CcO uses the energy released on reduction of O2, the product of the OEC reaction, to add to the electrochemical proton gradient. The problems of current interest are somewhat different in the two proteins. In PSII the open questions focus on the details of the S-state cycle, asking how the OEC structure, electron distribution,
and protonation states prepare the system for the O–O bond formation. There are several well defined water channels which are assumed to be always open for their business of transporting water, \( \text{O}_2 \), and protons. In contrast, in CcO it is the identification of the proton pathways and the method by which they are controlled that poses the most challenging questions. A variety of computational approaches—including QM/MM, MD, and CE/MC have been applied to address these questions. Each method has its strengths and weaknesses, but in combination they are producing a picture of the detailed mechanisms of these two important proteins.

**Author Contributions:** D.K. and X.C. contributed to the origin and draft preparation focused on PSII and CcO respectively. U.K., Y.Z., J.M., and M.M. added to the review of methodology. M.R.G. Gunner reviewed and edited the manuscript.

**Funding:** This research was funded by the Department of Energy Basic Energy Sciences grant number DE-SC0001423 and the National Science Foundation grant MCB-1519640.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

7. Styring, S.; Sjöholm, J.; Mamedov, F. Two tyrosines that changed the world: Interfacing the oxidizing power of photochemistry to water splitting in photosystem II. *Biochim. Biophys. Acta* 2012, 1817, 76–87. [CrossRef]


31. Retegan, M.; Krewald, V.; Mamedov, F.; Neese, F.; Lubitz, W.; Cox, N.; Pantazis, D.A. A five-coordinate Mn(IV) intermediate in biological water oxidation: Spectroscopic signature and a pivot mechanism for water binding. *Chem. Sci.* 2015, 7, 72–84. [CrossRef]


37. Boussac, A.; Rutherford, A.W. Comparative study of the g=4.1 EPR signals in the S2 state of photosystem II. *Biochim. Biophys. Acta* 2000, 1457, 145–156. [CrossRef]


41. Wang, J.; Askerga, M.; Brudvig, G.W.; Batista, V.S. Crystallographic Data Support the Carousel Mechanism of Water Supply to the Oxygen-Evolving Complex of Photosystem II. *ACS Energy Lett.* 2017, 2, 2299–2306. [CrossRef]


123. Ho, F.M.; Styring, S. Access channels and methanol binding site to the CaMnOX cluster in photosystem II based on solvent accessibility simulations, with implications for substrate water access. *Biochim. Biophys. Acta* 2008, 1777, 140–153. [CrossRef]
143. Dilbeck, P.L.; Bao, H.; Neveu, C.L.; Burnap, R.L. Perturbing the water cavity surrounding the manganese cluster by mutating the residue D1-Valine 185 has a strong effect on the water oxidation mechanism of photosystem II. *Biochemistry 2013*, *52*, 6824–6833. [CrossRef]
144. Debus, R.J. Evidence from FTIR difference spectroscopy that D1-Asp61 influences the water reactions of the oxygen-evolving Mn$_4$CaO$_5$ cluster of photosystem II. *Biochemistry 2014*, *53*, 2941–2955. [CrossRef] [PubMed]


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).