Photosynthesis Research (2023) 156:101–112
https://doi.org/10.1007/s11120-022-00973-0

ORIGINAL ARTICLE

Tools for analyzing protonation states and for tracing proton transfer pathways with examples from the *Rb. sphaeroides* photosynthetic reaction centers

Rongmei Judy Wei[1,2] · Umesh Khaniya[2,3] · Junjun Mao[2] · Jinchan Liu[4] · Victor S. Batista[5] · M. R. Gunner[1,2,3]

Received: 2 July 2022 / Accepted: 3 October 2022 / Published online: 29 October 2022

Abstract

Protons participate in many reactions. In proteins, protons need paths to move in and out of buried active sites. The vectorial movement of protons coupled to electron transfer reactions establishes the transmembrane electrochemical gradient used for many reactions, including ATP synthesis. Protons move through hydrogen bonded chains of waters and hydroxy side chains via the Grotthuss mechanism and by proton binding and release from acidic and basic residues. MCCE analysis shows that proteins exist in a large number of protonation states. Knowledge of the equilibrium ensemble can provide a rational basis for setting protonation states in simulations that fix them, such as molecular dynamics (MD). The proton path into the Q_B site in the bacterial reaction centers (RCs) of *Rb. sphaeroides* is analyzed by MD to provide an example of the benefits of using protonation states found by the MCCE program. A tangled web of side chains and waters link the cytoplasm to Q_B. MCCE analysis of snapshots from multiple trajectories shows that changing the input protonation state of a residue in MD biases the trajectory shifting the proton affinity of that residue. However, the proton affinity of some residues is more sensitive to the input structure. The proton transfer networks derived from different trajectories are quite robust. There are some changes in connectivity that are largely restricted to the specific residues whose protonation state is changed. Trajectories with $Q_B^{\bullet-}$ are compared with earlier results obtained with Q_B [Wei et. al Photosynthesis Research volume 152, pages153–165 (2022)] showing only modest changes. While introducing new methods the study highlights the difficulty of establishing the connections between protein conformation.

Keywords Semiquinone · Protonation states · Microstates · Proton transfer · Reaction center · Hydrogen bond network

Introduction

The light reactions of photosynthesis occur in proteins embedded in bacterial or chloroplast membranes. These proteins bind cofactors that carry out light absorption and the core cofactors participate in electron and proton transfers (Gunner et al. 2013). The reactions are initiated by absorption of light by chlorophyll, carotenoids and other chromophores followed by electron transfer across the proteins. The directionality of electron transfer and the coupled uptake and release of protons generates the transmembrane electrochemical potential gradient (Gunner et al. 2013). These reactions polarize the membrane creating a positive potential in the periplasm of bacteria and lumen of chloroplasts with the cytoplasm or stroma being more negative. Protons accumulated on the lumen side are transferred down the electrochemical gradient in the F1/F0 ATP synthase to drive the synthesis of ATP. The electron transfer chains in green

✉ M. R. Gunner
mgunner@ccny.cuny.edu; marilyn.gunner@gmail.com

1 Ph.D. Program in Chemistry, The Graduate Center, City University of New York, New York, NY 10016, USA

2 Department of Physics, City College of New York, New York, NY 10031, USA

3 Ph.D. Program in Physics, The Graduate Center of the City University of New York, New York, NY 10016, USA

4 Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA

5 Department of Chemistry, Yale University, New Haven, CT 06520, USA

plants lead to high energy reduced products such as NADH which in conjunction with ATP reduces CO_2 to sugar. In contrast, in the purple non-sulfur bacteria described here the type II photosynthetic reaction centers (RCs) work in an electron transfer cycle with the cytochrome bc_1 complex. The pair of proteins use cytochrome c and ubiquinone as mobile electron carriers to simply increase the proton gradient (Blankenship 2021).

The reaction centers (RCs) from the purple non-sulfur bacterial *Rb. sphaeroides* are analyzed to describe methods that characterize the distribution of protonation states and networks for proton transfer, reviewing and extending work from earlier studies (Khaniya et al. 2022; Wei et al. 2022). Each excitation of the protein leads to reduction of a ubiquinone at the Q_A site which in turn reduces the secondary quinone, Q_B. First a semiquinone $Q_B^{\bullet-}$ is formed, followed by protonation, preceding the second reduction and second protonation (Paddock et al. 1999). QH_2 then is released from the protein into the membrane and replaced with an oxidized quinone. Experimental and computational studies have explored the electron and proton transfers to Q_B (Okamura et al. 2000; Wraight 2004), the thermodynamics of the Q_A and Q_B reduction and protonation reactions (Gunner et al. 2008) and the proton transfer pathways from the cytoplasm to the Q_B site (Ishikita and Knapp 2004; Krammer et al. 2009; Wei et al. 2022). The Q_B site is ≈15 Å from the cytoplasm requiring long-range transfer to bring protons to the reduced quinone. Q_B is surrounded by a very complex web of polar and protonatable residues, which modulate the quinone redox potential and serve as proton transfer paths (Sebban et al. 1995).

Proton transfer networks are established with water, His and hydroxyl containing residues, including Ser, Thr, Tyr involved in Grotthuss-type mechanisms. In addition, acidic and basic residues serve as proton loading sites that release and bind protons translocated along hydrogen bonds (Ishikita and Saito 2014; Bondar 2022b). Surrounding non-Grotthuss competent residues such as Asn, Gln, Trp can orient and stabilize the connections needed for proton conduction.

Proton transfer pathways are found with different motifs. They can be linear chains with a single amino acid at the entry and exit and few branches along the path. Examples include the pathways starting approximately 10 Å from the oxygen evolving complex (OEC) in photosystem II (PSII) (Vassiliev et al. 2012; Kaur et al. 2019; Hussein et al. 2021), the D and K pathways in cytochrome c oxidase (CcO) (Sharpe and Ferguson-Miller 2008; Kaila et al. 2010) the three antiporter derived paths in Complex I (Sazanov 2015; Kaila 2018) and in bacteriorhodopsin (Balashov 2000). These networks often rely on proton transfer through key amino acid residues. There are also very tangled paths as found here near Q_B in RCs (Wei et al. 2022), on the proton exit side in CcO (Cai et al. 2018), the E-channel of Complex I (Kaila 2018; Khaniya et al. 2020) and the region within 7–10 Å of the OEC in PSII (Kaur et al. 2021b). We do not know why proteins choose to use simple or complex motifs. Single mutations in linear paths often stop proton transfer (Kaila et al. 2010) while complex paths often require multiple mutations to block proton transfer (Okamura et al. 2000; Wraight 2004). Linear paths can often be traced by visual inspection of a protein structure file (Sharpe and Ferguson-Miller 2008), but complex motifs must be characterized by computational techniques such as those described here.

Proton transfers accompany electron transfers to reduce the change in charge induced by the redox reactions (Merchant and Sawaya 2005). Quinones and other cofactors such as the Mn_4O_5Ca OEC of PSII will change protonation through the reaction cycle (Okamura et al. 2000; Wraight 2004; Gunner et al. 2008; Vinyard and Brudvig 2017). Other redox cofactors such as chlorophylls, iron sulfur clusters and hemes (Reedy and Gibney 2004) do not change protonation state. However, when a cofactor changes charge the protonation states of nearby residues often change (Wraight 2004). A number of computational techniques have been developed to follow the changes in residue protonation states coupled to quinone reduction (Beroza et al. 1991; Rabenstein et al. 1998; Alexov and Gunner 1999; Ullmann and Knapp 1999; Ishikita and Knapp 2004). Thus, the first reduction of Q_B leads to proton uptake to residues near the quinone that shift their proton affinity in the presence of the introduced negative charge. In contrast, the second reduction requires proton binding to the quinone itself.

Approximately 25% of the amino acid residues in an average protein structure are acidic or basic residues, Asp, Glu, Arg, Lys and His (Kim et al. 2005). The protonation state of acids and bases and the redox state of cofactors as well as the orientation of polar side chain and amide backbone dipoles generate electrostatic fields that regulate the structure and function of the protein (Gunner et al. 1996; Fried and Boxer 2015). In light harvesting proteins, the fields tune the spectra so that the absorbance of chemically identical chromophores can span a significant energy range (Friedl et al. 2022). Likewise, the cofactor redox potentials (E_ms) (Ullmann et al. 2002; Mao et al. 2003; Klingen and Ullmann 2004; Reedy and Gibney 2004; Zheng and Gunner 2009) and site pK_as (Pahari et al. 2019) are determined by the electrostatic potentials. The protein environment thus allows the chemically identical chromophores, redox cofactors and amino acids to have significantly different absorption spectra and electron and proton affinities in their individual sites. Without this tuning, proteins would likely need many more types of ligands to achieve a comparable range of functionality.

Simulations of proton and redox titrations generally provide the average protonation and redox state at a given pH and E_h usually determined by Monte Carlo (MC) sampling

(Alexov et al. 2011; Gunner and Baker 2016). One class of programs allow limited sampling of the conformations in the protein (Gunner and Baker 2016). For example, the program MCCE (Multi-Conformation Continuum Electrostatics), which will be used here allows for side chain rotamer sampling on a fixed backbone. This class of programs efficiently come to equilibrium in almost all cases within the physically constrained protein conformation. Although it should be noted that when many groups are coupled together, such as the waters in a wire, convergence can be difficult to achieve (Zhang et al. 2021). One work around for the rigid backbone is to carry out multiple independent calculations of protonation states and side chain conformations starting with different experimental structures (Lu and Gunner 2014) or with multiple MD frames (Cai et al. 2018; Cai et al. 2020).

An evolving approach, constant pH MD (cpHMD), is designed to meld protonation changes with MD, often carrying out a MC protonation state analysis at fixed intervals within the MD trajectory (Baptista et al. 1999, 2002; Teixeira et al. 2002; Damjanovic et al. 2018; Liguori et al. 2019; Sarkar et al. 2019; Vila-Vicosa et al. 2019; Cruzeiro et al. 2020; Torabifard et al. 2020; Vilhjalmsdottir et al. 2020; da Rocha et al. 2022). Following each change in charge state, the system must be relaxed. pH replica exchange methods help to speed convergence (Itoh et al. 2011). Currently cpHMD converges efficiently for systems with a few titrating sites, but it is challenging to converge with large proteins with many, interacting buried charges. Studies of larger systems profit from recognizing the residues that are firmly ionized or neutral and removing them from consideration (Baptista et al. 2002). Thus, compromises made in programs such as MCCE and by cpHMD show the difficulty of bringing the protonation and conformational states into equilibrium as is needed to explore proton transfer processes.

In contrast, standard Molecular Dynamics (MD) simulations focus on exploring conformational space, picking one protonation state for each residue that will be fixed for the whole trajectory. Likewise, the simulations that include quantum effects such as DFT and QM/MM simulations have a fixed number of protons. Protons can shift between neighbors in the portions of calculations that incorporate quantum mechanics, but the number of protons remains the same.

Monte Carlo (MC) sampling is an ideal approach to find the Boltzmann distribution of protonation states as protons bound to the acidic and basic residues in the protein come to equilibrium with those in a bath at a fixed chemical potential (i.e., pH) (Baptista et al. 1999). MC sampling works by randomly generating microstates, where each is a single choice for each residue and ligand given their degrees of freedom. For example, in MCCE a "microstate" has a defined position for all side chains, protonation state of each acidic and basic residue, cofactor redox state and ligand position (Song et al. 2009). The energy of each new microstate is compared to that of the previous one and Metropolis–Hastings sampling decides if the new microstate is accepted or if the previous one is retained. Recent work has begun to analyze the distribution and properties of the accepted MC microstates generated by MCCE (Khaniya et al. 2022). The MCCE ensemble of microstates will be used here to find the distribution of "protonation microstates", each of which define the protonation state of each acidic and basic residue independent of the side chain conformations (Khaniya et al. 2022).

The examination of protonation microstates shows that proteins are not in a single protonation state. Depending on the entropy contributed by the conformational degrees of freedom, the protonation microstates that are most probable may not be at the lowest energy. Different microstates generate different electrostatic potentials through the protein, thus will lead to spectral width, and the modulation of E_ms and pK_as of buried groups. Analysis of the residue protonation states in all accepted MC microstates can identify complex connections showing which residues are coupled to each other. Correlation of protonation and redox states can show how the ensemble of protonation states shift to relax in redox reactions (Baptista et al. 1999; Teixeira et al. 2002; Mao et al. 2003; Khaniya et al. 2022). A simple yet important use of the information about individual protonation microstates is that they provide a rational choice for all protonation states for input to MD or other higher-level calculations.

Tracing hydrogen bonded paths requires knowing the position of neighboring residues and waters in individual conformational microstates. The hydrogen bonds found in the many accepted MCCE microstates can be built into proton transfer chains using network analysis (Cai et al. 2018; Khaniya et al. 2020; Wei et al. 2022). Connections can also be determined in the many frames (snapshots) of an MD trajectory (Bondar et al. 2010). When hydrogen bonds are found in an MD trajectory the protonation state of acidic and basic groups are fixed which determines whether they can be a proton donor or acceptor in the path. In MCCE these groups are allowed to sample different protonation states so can switch the role they play in the proton transfer path.

The results provided here will use the classical MD package OpenMM (Eastman et al. 2013, 2017) and the MCCE MC sampling program (Song et al. 2009). Similar analysis has been carried out on bacterial RCs (Wei et al. 2022), PSII (Kaur et al. 2021b), cytochrome c oxidase (Cai et al. 2020) and Complex I (Khaniya et al. 2020) and reviewed in (Kaur et al. 2019, 2021a). We will use MCCE protonation microstates, which define the protonation state of every acid and base, to initiate MD. The proton transfer pathways will be obtained from the hydrogen bonds in the MD trajectory. The effects of the MD protonation states on the hydrogen bonded paths in the

MD trajectories and on the MCCE calculated proton affinities will be described.

Methods

Choice of protonation states in the MD trajectories

The coordinates of the *Rb. sphaeroides* photosynthetic reaction centers (RCs) (PDB ID: 1AIG) with Q_B in the proximal site are the starting material for the MCCE and MD simulations (Stowell et al. 1997). MCCE uses Monte Carlo sampling to find the Boltzmann distribution of protonation states. Several highly probable protonation states that differ in their protonation of residues near Q_B are chosen as input for MD simulations.

The MCCE calculations use a protein dielectric constant of 4, and 80 for water both outside and in cavities within the protein. Only isosteric conformer sampling is used, which allows a search of the positions of polar protons and His tautomers and Asn and Gln rotamers. During the MCCE sampling, Q_A and the other redox cofactors are always fixed in their neutral ground state. In contrast, Q_B is fixed in the anionic ($Q_B^{\bullet-}$) semiquinone form. The MCCE calculations (Song et al. 2009) are carried at pH 7 as described in previous work (Khaniya et al. 2022). After the calculation, MCCE finds most Asp, Glu, Lys and all Arg are fixed ionized and most His are neutral with the proton on ND1 (HSD). These choices are maintained in the MD simulations. In line with the results of the MCCE calculations, Glu H43, L104 and Lys H197 are always fixed neutral and His H126, 128 and 301 are always fixed ionized. HisH98, HisH204, HisL116, HisL168, HisL211, HisM145 are fixed with the proton on NE2 (HSE). Then the protonation states of six other residues are modified in each trajectory. AspL210, AspL213, GluL212 and HisH68 switch their protonation states between the trajectory with the neutral Q_B (Q1) (Wei et al. 2022) and $Q_B^{\bullet-}$ (SQ1-SQ4), shown in Table 1. In the trajectories SQ2, SQ3 and SQ4, the total charge is maintained with a proton shifting between residues. LysH146 is not in the group of interconnected residues around Q_B, but it may affect the protonation states of nearby residues.

Table 1 Protonation states of residues near Q_B that are modified in the different MD trajectories

	Trajectory				
	Q1	SQ1	SQ2	SQ3	SQ4
Q_B	Q_B	$Q_B^{\bullet-}$	$Q_B^{\bullet-}$	$Q_B^{\bullet-}$	$Q_B^{\bullet-}$
ASP L210	0	− 1	− 1	− 1	− 1
ASP L213	− 1	0	0	− 1	0
GLU L212	− 1	0	0	0	0
GLU M236	0	0	− 1	0	0
HIS H68	0	1	1	1	0
LYS H146	1	0	0	0	0
LYS H0197	1	0	0	0	0
Sum H^+	0	0	− 1	− 1	− 1
Total H^+	3	3	2	2	2

The MD parameters and settings

We prepared the topology and parameter files for molecular dynamics using the CHARMM GUI (Jo et al. 2008) with the protein cofactors patched into the protein, as described in our previous work (Wei et al. 2022). Cofactor parameters are from Matyushov (Ceccarelli and Marchi 2003; LeBard and Matyushov 2008) and can be downloaded from https://github.com/GunnerLab/RCsparameters. New parameters were created for the ubisemiquinone ($Q^{\bullet-}$) using the VMD toolkit (Mayne et al. 2015) as described in the SI.Part I. The $Q^{\bullet-}$ parameters can be found at https://github.com/judywei2333/semi_brc_parameters.

MD is carried out in a POPC membrane with explicit water in a rectangular box of 115 Å × 115 Å × 120 Å using OpenMM (Eastman et al. 2013, 2017); 86 Na^+ and 84 Cl^- are included for the charge neutrality. Each MD production run is 200 ns long for each protonation state with $Q_B^{\bullet-}$ and compared with previous results with the neutral Q_B (Wei et al. 2022). The trajectory coordinates are saved every 10 ps and is written out in a dcd file. The snapshots spaced 20 ns apart are used for the analysis here.

Determination of the correlation between residue protonation in the protonation microstate ensembles

The protonation microstates are found in MCCE. Each unique protonation microstate has a different distribution of protons over the acidic and basic residues. Individual protonation microstates may have the same number of protons bound to different residues (protein tautomers) or different numbers of protons and thus a different charge. The weighted Pearson's correlation is applied here to investigate the charge correlations between residues in the millions of generated microstates (Khaniya et al. 2022), although there are other methods to find coupling between residues in the large dataset (Baptista et al. 1999). All microstates with the same protonation distribution are identified as a unique microstate and each is associated with a count of the number of times this protonation microstate is found in the accepted MC ensemble. Using the count of each unique microstate as a weight is equivalent to the standard Pearson's correlation

where every accepted microstate in the sampling sequence is counted once. The weighted Pearson's correlation coefficient is:

$$r_{pq} = \frac{\sum_{i=1}^{n} w_i(p_i - \overline{p})(q_i - \overline{q})}{\sqrt{\sum_{i=1}^{n}(w_i(p_i - \overline{p})^2)\sum_{i=1}^{n}(w_i(q_i - \overline{q})^2)}} \quad (1)$$

The count of each individual conformation/protonation microstate is the number of times it is chosen in MC sampling. A microstate count increases when the subsequent randomly chosen microstate is rejected in Metropolis–Hastings sampling. The protonation microstate count is the sum of the counts for the many conformational microstates with the same number and position of protons but different side chain positions.

The correlation is determined given each protonation microstate i, and the charges of the two residues to be compared: p_i and q_i. The total number of unique accepted protonation microstates is n. The average charge of these residues in the full MC ensemble is $\overline{p}$ and $\overline{q}$. The weight here is equal to the count of protonation microstate i. The weighted coefficient is used to ensure that the correlation matrix is not dominated by the many unique protonation microstates with low probability in the ensemble. The weighted Pearson's correlation coefficient ranges from − 1 to 1, with 1 or − 1 indicating the sites are fully correlated or anticorrelated and 0 indicating that they are not correlated at all (or corresponding to orthogonal displacements).

Determination of the proton transfer paths in the MD trajectories

The goal is to find the hydrogen bonded paths that protons can take from residues at the cytoplasmic surface to $Q_B^{\bullet-}$ (Okamura et al. 2000; Wraight 2004). 2500 snapshots are selected from the 51–100 ns portion of the MD trajectory. The membrane and explicit waters with less than 20% surface accessibility are removed. Hydrogen bonds are identified by a distance between the hydrogen and a heteroatom between 1.2 and 3.5 Å and angle between the donor hydrogen and acceptor, that is greater than 90° (Wei et al. 2022). The hydrogen bond network is developed based on the depth-first search (DFS) to carry out an exhaustive search of the hydrogen bonds from the periplasm side of the RCs to the Q_B site (Mercado et al. 2022). The hydrogen bond network are visualized using Cytoscape (Shannon et al. 2003). Networks described here are mapped between residues with a maximum of two intervening waters. This leads to a complete map of the full hydrogen bond network. The networks with 0–4 intervening waters are provided in Fig SI.5. The network becomes more tangled when longer paths through more waters are considered.

Results

Distribution of protonation microstates

MCCE maintains the backbone of the input structure, so it cannot simulate displacements of distant sites or changes in the local structure that are backbone dependent. Standard MD in contrast can show the changes in the protein driven, in part, by one distribution of ionized residues. Here, five protonation microstates were chosen as input for MD trajectories that represent different states of six residues near Q_B whose protonation states were shown to be correlated with each other (Khaniya et al. 2022) (Table 1).

There are 132 protonatable residues in RCs whose Boltzmann distribution of protonation microstates are determined by MCCE using individual snapshots from the five trajectories. Figure 1A shows the distribution of protonation microstates derived from one snapshot from the MD trajectory SQ2 (see Table 1 for trajectory protonation). Here 23 residues have different protonation states in the accepted MC microstates. Each dot is one of the 5733 unique accepted protonation microstates. The graph is plotted with the natural log of unique microstate probability against the charge of the protein in that state. The dots in each column have the same net charge, ranging from − 2 to 7. The most occupied microstates have charges between 1 and 3. The Boltzmann averaged charge of the ensemble of acids and bases is 1.23 with a charge on $Q_B^{\bullet-}$ of − 1.

Many conformational microstates, each of which have different energy, are in the same protonation microstate. The dot size and color represent the range of energies found in the conformational ensemble for that pronation microstate. The range corresponds to the maximum minus the minimum energy for the microstates with this proton distribution. The most populated protonation microstates are found in a diverse group of conformational states with a wide range of energies.

The large number of protonation microstates contains many rare states while a small subset makes up the bulk of the population. Figure 1B shows that the protonation microstate with the highest probability contributes 8% of the population while 26 microstates make up about 50% of the ensemble. However, > 630 are needed to capture 90% of the total population. To tame the explosion of protonation microstates, the protein can be divided into clusters which are spatially separated so the charge of each cluster is relatively independent. A Q_B cluster can be defined as an extended group of residues in the Q_B proton transfer network

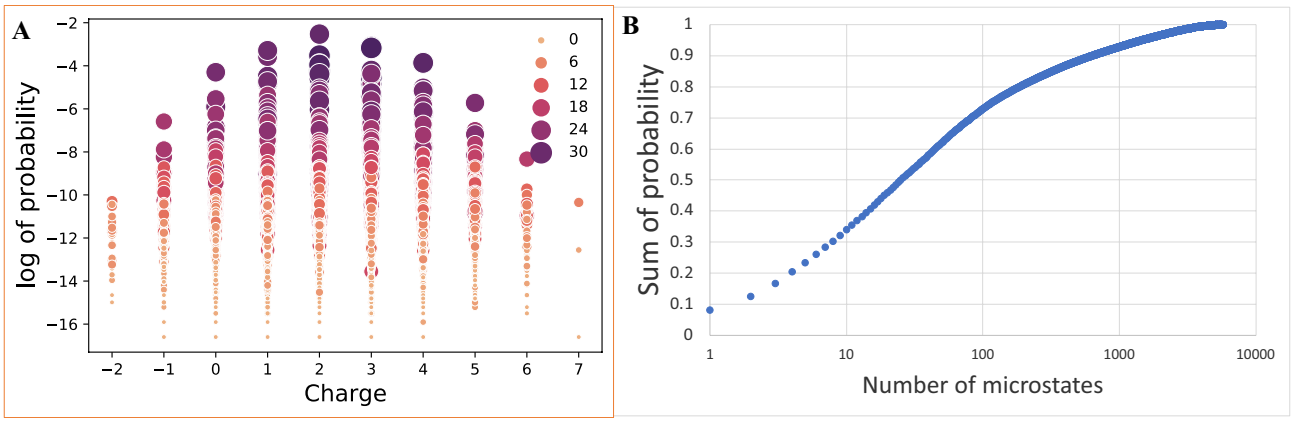

Fig. 1 **A** Distribution of unique RC protonation microstates at pH 7. The MCCE calculations were run on one snapshot of the MD trajectory SQ2 (Table 1). Each dot represents one protonation microstate, and natural log of the probability gives the fraction of times a state with this protonation distribution is accepted by MC sampling. Each protonation microstate can be found in many MCCE side chain conformation states with a range of microstates with different energy. The dot color and size report on this range of energies with bigger, darker dots having a wider energy distribution. Energies are in kcal/mol. **B** The running sum of the probabilities of the protonation microstates. Protonation microstates are graphed in order of decreasing probability. Each point adds the probability to that of all points to its left

(see below in Table 3). The cluster has 15 protonatable residues and 12 of them change protonation states in the MC sampling, leading to a total of 197 protonation microstates. Now only three protonation microstates constitute 50% of the population and 15 represent 90%. The 15 protonation microstates with higher probability for the Q_B cluster are reported in Table SI.1.

Having access to all protonation microstates allows us to find the correlation between different properties that are sampled in the calculations. Here, we show the weighted Pearson's correlation coefficient (Eq. 1) for the coupled protonation state of individual residues near $Q_B^{\bullet-}$ using a snapshot from the SQ2 trajectory (Fig. 2). The protonation states of three acidic residues are strongly correlated. GluL213 lies between AspL210 and GluL212 leading towards Q_B.

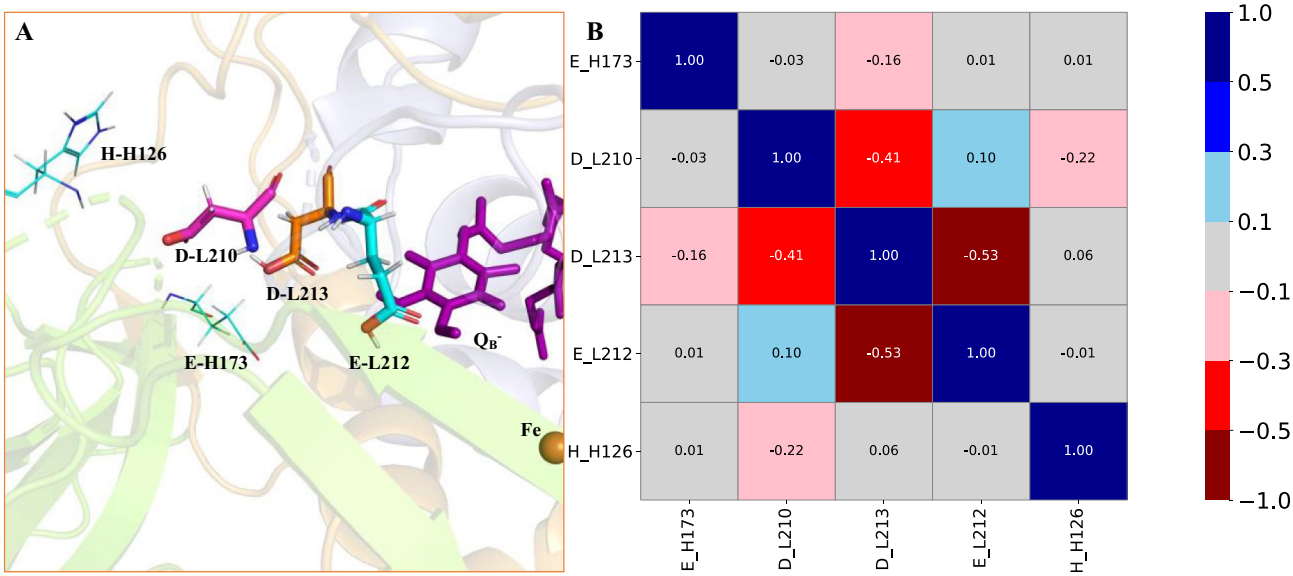

Fig. 2 The correlation of residue charge in the protonation microstates in the MCCE calculations with $Q_B^{\bullet-}$ of a snapshot from the MD trajectory SQ2. **A** Residues with significant correlation. The residues names are given as Chain Designator (H, L or M) and the one letter code for the amino acids followed by the residue number. Residues with stronger correlation in sticks, more weakly interacting residues are lines. $Q_B^{\bullet-}$: the ubiquinone head and tail as magenta sticks, Fe: yellow sphere. **B** weighted Pearson's correlation of protonation states of individual residues. Blue: positive correlation; red: negative correlation; darker color indicates stronger correlation

The protonation state of the central GluL213 is negatively correlated with the other two acids. This leads to a positive correlation between AspL210 and GluL212 so that both tend to be ionized, or to be neutral responding to the charge on AspL213. HisH126 and GluH173 both have significant correlation within this cluster. Thus, the cluster protonation microstates show the residues are coupled together. However, no correlation coefficient is close to ± 1, indicating that residues do not change in lockstep with each other. This makes it challenging to discover these connections without using statistical tools for evaluating large ensembles (Baptista et al. 1999; Khaniya et al. 2022).

The impact of fixed MD protonation states on the proton affinity found in MCCE

We can consider how the conformation input from MD configurations affect the proton affinity. MCCE simulations were run on 12 snapshots, four are from a trajectory with Q_B and two each from four different trajectories with $Q_B^{\bullet-}$. The results for each residue were separated by whether the trajectory had this residue fixed to be ionized or neutral in MD (Table 2). In all cases MCCE favors the charge state imposed by the MD trajectory for a given residue. The variation in average protonation can be modest, as for HisH68, or more substantial as for AspL213 or GluM236. It should be noted that, with the exception of the stable value for HisH68, the standard deviation is largest when the average protonation is near 50% ionized. This is near the steepest part of a titration curve, where smaller changes in energy lead to larger changes in proton binding.

A more detailed version of Table 2 is given in SI. 2, which provides the MCCE calculated protonation states in each snapshot for the 7 residues whose protonation is modified in the different trajectories as well as that of 10 other residues near Q_B whose charge is the same in all MD trajectory. Between 2 and 4 snapshots were chosen from each trajectory. In general, there is agreement in the MCCE average protonation in the snapshots from the same trajectory. However, this is not always the case. GluM236 ranges from being fully neutral to 85% ionized in different snapshots from the Q1 trajectory where this Glu is fixed neutral. LysH197 also gives quite different ionization in different snapshots from two of the trajectories. The residues of interest were chosen as being highly correlated with other residues in the initial MCCE calculations of the crystal structure. We find here that their protonation can sometimes be influenced by details of the structure in addition to their protonation state in MD.

The protonation states of other residues near Q_B are listed in Table SI.2C. Their average protonation does not change much between different snapshots. The exception is HisH128 which is at the entrance to a proton input channel.

Thus, the protonation in the trajectory generally leads to a structure that stabilizes the charge on that residue favoring the protonation state used in MD. However, comparisons of the average proton affinity from individual snapshots show that the protonation states of some residues are significantly influenced by the local environment along the MD trajectory. These may profit from analysis by cpHMD. Another fruitful approach to this real problem is to subject the trajectory to meta-analysis focusing on the positions of residues with highly variable proton affinity. Previous calculations have shown that it can be possible to identify geometric arrangements that switch the proton affinity of a residues between different states (Cai et al. 2020).

Finding proton transfer paths

Hydrogen bonded networks that connect the cytoplasm to Q_B are found in the RC MD trajectories. As with previous calculations of hydrogen bonded paths in RCs (Wei et al. 2022), the hydrogen bond networks find the residues connected by waters. Proton transfer networks found from the MD trajectory can bring protons from the negative, cytoplasmic side of the membrane to Q_B via the Grotthuss transfer mechanism augmented by transient proton trapping on acidic and basic residues. Here, residues are viewed as connected if a path between them can be found via 2 or fewer

Table 2 MCCE calculated average charge state for residues that are fixed in different protonation states in MD (Table 1)

	Neutral			Ionized		
	average	stdev	N	average	stdev	N
Asp L210	− 0.01	0.02	4	− 0.93	0.11	8
Asp L213	− 0.27	0.30	6	− 0.98	0.02	6
Glu L212	− 0.16	0.18	8	− 0.52	0.48	4
Glu M236	− 0.44	0.40	10	− 1.00	0.01	2
His H68	0.44	0.23	6	0.74	0.20	6
Lys H197	0.77	0.37	8	1.00	0.00	4
Lys H146	1.00	0.01	8	1.00	0.00	4

'Neutral' gives the average protonation in snapshots coming from trajectories where the residue is neutral, while 'Ionized' indicates the residue was ionized in the trajectory. N is the number of snapshots averaged

Table 3 Count of the number of connections each residue makes in the network with a maximum of two intervening waters

(A)	Q1	SQ1	SQ2	SQ3	SQ4	Ionized	Neutral
QB	2	3*	3*	4*	3*	3.25	2
Asp L210	8	7*	7*	8*	6*	7.00	8
Asp L213	5*	5	6	6*	5	5.5	5.33
Glu L212	4*	2	3	2	2	4	2.25
Glu M236	5	3	4*	4	3	4	3.75
His H68^	3	3*	9*	3*	2	4.25	2.50

(B)	Q1	SQ1	SQ2	SQ3	SQ4	av connect	±	stdev
*His H126^	10	9	16	12	8	11.0	±	3.16
Thr L226	7	7	7	7	8	7.2	±	0.45
*Lys H130	9	5	7	9	5	7.0	±	2.00
Ser L223	9	6	7	4	7	6.6	±	1.82
*Glu H173	6	5	8	5	4	5.6	±	1.52
*Asp M17	4	4	7	6	4	5.0	±	1.41
*Glu H224^	5	5	5	5	5	5.0	±	0.00
Thr L214	6	5	7	4	3	5.0	±	1.58
*Asp H124	7	4	7	3	3	4.8	±	2.05
*His H128^	1	6	6	6	3	4.4	±	2.30
Tyr M3^	1	5	5	6	4	4.2	±	1.92
*Lys H132^	3	4	4	5	3	3.8	±	0.84
Thr L208	6	0	6	5	1	3.6	±	2.88
*Asp M240	1	1	4	4	2	2.4	±	1.52
Thr M21^	0	0	3	4	3	2.0	±	1.87
Ser M8	0	0	3	3	3	1.8	±	1.64
His L190	1	1	1	1	2	1.2	±	0.45

(C)	Q1	SQ1	SQ2	SQ3	SQ4	av crg	±	stdev
Q_A	2	2	2	3	2	2.2	±	0.45
Thr M222	1	1	1	1	1	1	±	0
His M219	1	1	1	1	1	1	±	0
Thr M261	0	0	0	2	0	0.4	±	0.89

(A) The residues with different charges in different trajectories. The count of connections is averaged separating the values from trajectories where the residue is ionized or neutral. (B) The number of residues connected to the residues with the same charge in all trajectories are sorted in the descending order of the average number of connections. (C) Connectivity of residues around Q_A

^ Surface residues

*Ionized residues

waters. There are 17 residues that are connected in at least one trajectory (Table 3A and B). A more complete list of 23 residues connected by four or fewer waters is provided in Table SI.3.

Table 3A shows the average number of connections made by the residues whose protonation state is different in the five trajectories. Residues tend to make more connections in trajectories where they are ionized than in the trajectories where they are neutral. Table 3B shows residues whose charge is the same in all trajectories, sorted by the average number of connections. The number and identity of the connections to other residues are similar in the different trajectories. The surface residue HisH126 can be connected to the most other residues. SerL223 and ThrL226, which are hydrogen bonded to Q_B also have many partners allowing protons to follow many paths to the quinone. The trajectory with Q_B neutral (Q1) was compared with those with Q_B ionized (SQ1–SQ4). There are only a few significant differences such as the surface residue TyrM3 being less connected in Q1 and GluM234 being more connected. As these are not near Q_B, this is likely to represent the vagaries of a given MD trajectory.

Table 3C shows the network around Q_A. ThrM222 and HisM219 make hydron bonds directly to the quinone. In the previous analysis of a trajectory where both quinones are neutral (Q1), Q_A never had any additional connections

beyond these two residues (Wei et al. 2022). Here, we see in SQ3 that Q_A can also connect to ThrM261, which connects to the $Q_B^{\bullet -}$ network. The role of Q_A is to be reduced only to the level of the semiquinone, passing this electron to Q_B and then to $Q_B^{\bullet}H$. $Q_A^{\bullet -}$ should never bind a proton and its isolation from a robust proton transfer network protects it from doing so.

Figure 3 shows a network aggregated from the 2500 snapshots from the 50–100 ns segment of the trajectory SQ2 The paths shown here simply trace connectivity, but do not show which paths are favored. Only residues that can participate in Grotthuss proton transfer or by binding and releasing protons are shown. In Fig. 3B each column represents one step away from Q_B. The nodes (residues)

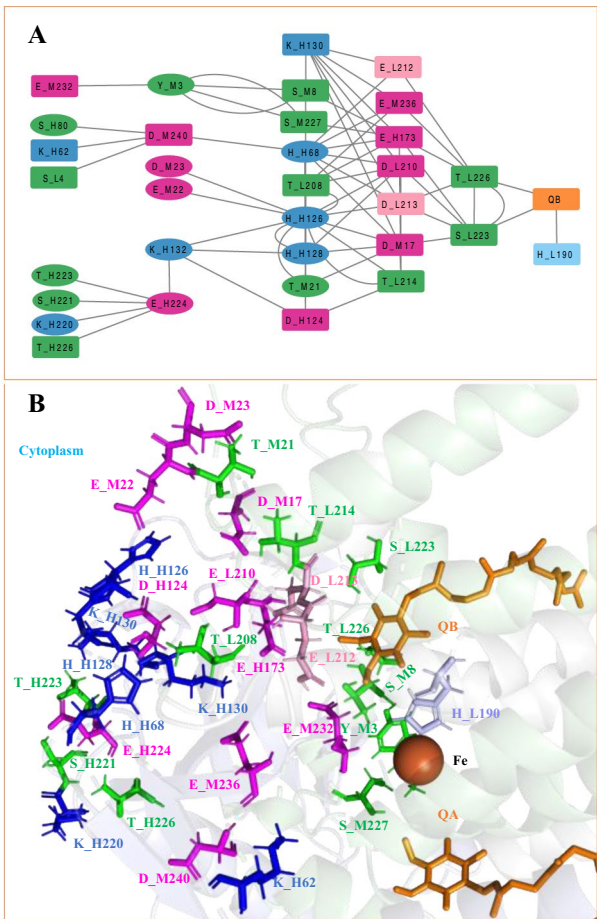

Fig. 3 Hydrogen bond network leading to $Q_B^{\bullet -}$ in a snapshot from SQ2. Residues are colored with: Q_B orange; Grotthuss competent hydroxyl residues greens; the bases blue, and acids red with darker shade indicating the residue is charged and the lighter shade that it is neutral. **A** Network connection drawn with Cytoscape (Shannon et al. 2003). Each node is a residue. Each line is a connection through two or fewer waters. **B** Pymol picture of the SQ2 snapshot from the network shown in (**A**). Residues are labeled with one letter amino acid code_subunitResidueNumber. Cytoplasm is on the left side as noted. S_H80 and S_L4 could not be shown in this orientation of the protein

can be connected by as many as two waters, which are not shown. Thus, SerL223, HisL190 and ThrL223 are the primary nodes connected to Q_B. followed by GluL212, M236, H173 and Asp L210 and L213 which are connected to Q_B through the Ser or Thr. The residues in the network are entangled with one another. Depending on the path a residue can serve as primary or secondary connections to Q_B. For example, SerL223 can connect directly to Q_B or connect via ThrL226.

The networks show connectivity although the preferred paths are yet to be determined (Ishikita and Saito 2014; Kaur et al. 2021a; Bondar 2022a). Table SI.4A shows the frequency of each of the 89 residue-residue connections found in at least one frame (snapshot) in the four SQ trajectories with at most two intervening waters. Focusing on SQ2 we see the Ser L223 is a hydrogen bond donor to Q_B 90% of the time in the MD trajectory, while the Thr L226 connects to either Q_B or to the Ser < 1% of the time. Thus, the proton is more likely to transfer via the Ser than the Thr. Ser L223 is often a hydrogen bond donor to Asp L213 and more rarely to GluH173 and AspM17 and occasionally to AspL210 and ThrL226. When protons transfer from the cytoplasm to Q_B the Ser would need to be a proton acceptor, with one of these acids protonated (Ishikita and Knapp 2004). Table SI.4B shows that overall, Ser and Thr are more likely to be proton donors than acceptors. This may be because there are more acids than bases in the network or because of the propensity of hydroxyl containing side chains to be better proton donors (Kim et al. 2005). Despite some acids being protonated here, none are found to satisfy the rules to be classified as donors in these trajectories.

Table 3A shows that individual residues tend to make connections to more other residues in the trajectories when they are ionized. However, Table SI.4A provides a mixed message about the changes in the persistence of the connections when groups are ionized. Thus, in SQ2 when both Glu M236 and His H68 are ionized, they are connected to each other in 61% of the frames and Glu M236 connects to Lys H130 72% of the time. However, in SQ4 when both His H68 and Glu M236 are neutral the frequency of their connection decreases to 14% and now Glu M236 connects to Lys H130 in only 12% of the frames. In another example, GluL213 loses its connection to ServL223 when it is ionized but strengthens its connection to Lys H130. However, overall, the frequency of connections to the six residues whose protonation states are changing are not strongly correlated with the residue charge. Thus, the network maintains many of the key connections, but the strength of connections can be dependent on the input residue charge state.

We find many proton transfer paths to Q_B^- (Fig. 3 and SI.5). There are four entries to the network on the surface

from HisH126, HisH128, HisH68 and TyrM3. All three His are protonated in the MD trajectories. The pathways with $Q_B^{\bullet-}$ are shorter than those found previously in trajectories with the neutral Q_B (Wei et al. 2022). Complete proton transfer paths can be found to each entry needing connections via no more than two waters. For example:

SQ2 entry 1:

$$HisH126^{+} \rightarrow AspM17^{-} \rightarrow 1w \rightarrow SerL223^{0} \rightarrow \mathbf{Q_B^{\cdot-}}$$

SQ2 entry 2:

$$HisH128^{+} \rightarrow 1w \rightarrow AspM17^{-} \rightarrow 1w \rightarrow SerL223^{0} \rightarrow \mathbf{Q_B^{\cdot-}}$$

SQ2 $Q_B\bullet$-:

$$HisH68^{+} \rightarrow 1W \rightarrow AspL210^{-} \rightarrow 2W \rightarrow SerL223^{0} \rightarrow QB^{\cdot-}$$

$Q_B\bullet$-

$$HisH68^{+} \rightarrow GluM236^{-} \rightarrow 2W \rightarrow ThrL226^{0} \rightarrow 1w \rightarrow QB^{\cdot-}$$

SQ2 entry4:

$$TyrM3^{0} \rightarrow 1W \rightarrow SerM8^{0} \rightarrow 2w \rightarrow GluH173^{-} \rightarrow 2w \rightarrow SerL223^{0} \rightarrow \mathbf{Q_B^{\cdot-}}$$

or

$$TyrM3^{0} \rightarrow 1W \rightarrow SerM8^{0} \rightarrow 2w \rightarrow GluH173^{-} \rightarrow 1w \rightarrow ThrL226^{0} \rightarrow 2W \rightarrow \mathbf{Q_B^{\cdot-}}$$

Conclusions

Here, we have explored the variation of protonation microstates in RCs. We found thousands of protonation states in the Boltzmann ensemble for the protein as a whole. When the protein is separated into smaller relatively independent regions, the combinatoric explosion can be significantly reduced. Thus, while more than 600 protonation microstates are needed to describe 90% of the proteins in the ensemble, only 15 protonation microstates contribute significantly to the ensemble for residues in the network near Q_B.

We have chosen different protonation states as inputs for MD trajectories. MD allows the protein to equilibrate and MCCE does find the proton affinity of each snapshot is dependent on the charge state in the input trajectory. However, while the proton affinity of many residues is fairly stable there are a number of residues where the MCCE output is quite sensitive to the input structure.

We have compared the proton transfer network in snapshots from multiple trajectories with different input protonation states. Overall, the proton transfer network is robust but when a group is charged it tends to make connections to more partners than when it is neutral. The persistence of the connections is more weakly dependent on the input protonation states.

Thus, it can be important to initiate MD with well-chosen protonation states (Kaila et al. 2014; Zheng and Cui 2017), such as those derived from MCCE microstate analysis which provides a self-consistent protonation state for all residues in the protein. If possible multiple high-probability protonation states can be used in different trajectories. Initial protonation states that are derived from the experimental structures may be less biased as the coordinates can represent an average of multiple protonation microstates in the crystal.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11120-022-00973-0.

Acknowledgements We acknowledge financial support from the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences, U.S. Department of Energy, Photosynthetic Systems grant DESC0001423. We thank the New York State's Graduate Research Training Initiative (GRTI) for support of a high-performance computing cluster at the City College of New York. We thank Joe Z Wu for help with the VMD molefacture setup. We thank Muhamed Amin, Gehan Ranepura and Jimin Wang for useful discussion.

References

Alexov E, Gunner MR (1999) Calculated protein and proton motions coupled to electron transfer: electron transfer from Q_A^- to Q_B in bacterial photosynthetic reaction centers. Biochemistry 38:8254–8270. https://doi.org/10.1021/bi982700a

Alexov E, Mehler EL, Baker N, Baptista AM, Huang Y, Milletti F, Nielsen JE, Farrell D, Carstensen T, Olsson MH, Shen JK, Warwicker J, Williams S, Word JM (2011) Progress in the prediction of pKa values in proteins. Proteins 79(12):3260–3275. https://doi.org/10.1002/prot.23189

Balashov SP (2000) Protonation reactions and their coupling in bacteriorhodopsin. Biochim Biophys Acta 1460(1):75–94. https://doi.org/10.1016/s0005-2728(00)00131-6

Baptista AM, Martel PJ, Soares CM (1999) Simulation of electron-proton coupling with a Monte Carlo method: application to cytochrome c3 using continuum electrostatics. Biophys J 76(6):2978–2998. https://doi.org/10.1016/s0006-3495(99)77452-7

Baptista AM, Teixeira VH, Soares CM (2002) Constant-pH molecular dynamics using stochastic titration. J Chem Phys 117(9):4184–4200. https://doi.org/10.1063/1.1497164

Beroza P, Fredkin DR, Okamura MY, Feher G (1991) Protonation of interacting residues in a protein by a Monte Carlo method: application to Lysozyme and the photosynthetic reaction center of *Rhodobacter sphaeroides*. Proc Natl Acad Sci USA 88:5804–5808. https://doi.org/10.1073/pnas.88.13.5804

Blankenship RE (2021) Molecular mechanisms of photosynthesis, 3rd edn. Wiley, Hoboken, NJ

Bondar AN (2022a) Graphs of hydrogen-bond networks to dissect protein conformational dynamics. J Phys Chem B 126(22):3973–3984. https://doi.org/10.1021/acs.jpcb.2c00200

Bondar AN (2022b) Mechanisms of long-distance allosteric couplings in proton-binding membrane transporters. Adv Protein Chem Struct Biol 128:199–239. https://doi.org/10.1016/bs.apcsb.2021.09.002

Bondar AN, Smith JC, Elstner M (2010) Mechanism of a proton pump analyzed with computer simulations. Theoret Chem Acc 125(3–6):353–363. https://doi.org/10.1007/S00214-009-0565-5

Cai X, Haider K, Lu J, Radic S, Son CY, Cui Q, Gunner MR (2018) Network analysis of a proposed exit pathway for protons to the P-side of cytochrome c oxidase. Biochim Biophys Acta Bioenerg 1859(10):997–1005. https://doi.org/10.1016/j.bbabio.2018.05.010

Cai X, Son CY, Mao J, Kaur D, Zhang Y, Khaniya U, Cui Q, Gunner MR (2020) Identifying the proton loading site cluster in the ba_3 cytochrome c oxidase that loads and traps protons. Biochim Biophys Acta Bioenerg 1861(10):148239. https://doi.org/10.1016/j.bbabio.2020.148239

Ceccarelli M, Marchi M (2003) Simulation and modeling of the Rhodobacter sphaeroides bacterial reaction center II: primary charge separation. J Phys Chem B 107(23):5630–5641

Cruzeiro VWD, Feliciano GT, Roitberg AE (2020) Exploring coupled redox and pH processes with a force-field-based approach: applications to five different systems. J Am Chem Soc 142(8):3823–3835. https://doi.org/10.1021/jacs.9b11433

da Rocha L, Baptista AM, Campos SRR (2022) Approach to study pH-dependent protein association using constant-pH molecular dynamics: application to the dimerization of β-Lactoglobulin. J Chem Theory Comput 18(3):1982–2001. https://doi.org/10.1021/acs.jctc.1c01187

Damjanovic A, Miller BT, Okur A, Brooks BR (2018) Reservoir pH replica exchange. J Chem Phys 149(7):072321. https://doi.org/10.1063/1.5027413

Eastman P, Friedrichs MS, Chodera JD, Radmer RJ, Bruns CM, Ku JP, Beauchamp KA, Lane TJ, Wang LP, Shukla D, Tye T, Houston M, Stich T, Klein C, Shirts MR, Pande VS (2013) OpenMM 4: a reusable, extensible, hardware independent library for high performance molecular simulation. J Chem Theory Comput 9(1):461–469. https://doi.org/10.1021/ct300857j

Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang LP, Simmonett AC, Harrigan MP, Stern CD, Wiewiora RP, Brooks BR, Pande VS (2017) OpenMM 7: rapid development of high performance algorithms for molecular dynamics. PLoS Comput Biol 13(7):e1005659. https://doi.org/10.1371/journal.pcbi.1005659

Fried SD, Boxer SG (2015) Measuring electric fields and noncovalent interactions using the vibrational stark effect. Acc Chem Res 48(4):998–1006. https://doi.org/10.1021/ar500464j

Friedl C, Fedorov DG, Renger T (2022) Towards a quantitative description of excitonic couplings in photosynthetic pigment-protein complexes: quantum chemistry driven multiscale approaches. Phys Chem Chem Phys 24(8):5014–5038. https://doi.org/10.1039/d1cp03566e

Gunner MR, Baker NA (2016) Continuum electrostatics approaches to calculating pK_as and Ems in proteins. Methods Enzymol 578:1–20. https://doi.org/10.1016/bs.mie.2016.05.052

Gunner MR, Nicholls A, Honig B (1996) Electrostatic potentials in *Rhodopseudomonas viridis* reaction center: implications for the driving force and directionality of electron transfer. J Phys Chem 100:4277–4291

Gunner MR, Madeo J, Zhu Z (2008) Modification of quinone electrochemistry by the proteins in the biological electron transfer chains: examples from photosynthetic reaction centers. J Bioenerg Biomembr 40(5):509–519. https://doi.org/10.1007/s10863-008-9179-1

Gunner MR, Amin M, Zhu X, Lu J (2013) Molecular mechanisms for generating transmembrane proton gradients. Biochim Biophys Acta 1827(8–9):892–913. https://doi.org/10.1016/j.bbabio.2013.03.001

Hussein R, Ibrahim M, Bhowmick A, Simon PS, Chatterjee R, Lassalle L, Doyle M, Bogacz I, Kim I-S, Cheah MH, Gul S, de Lichtenberg C, Chernev P, Pham CC, Young ID, Carbajo S, Fuller FD, Alonso-Mori R, Batyuk A, Sutherlin KD, Brewster AS, Bolotovsky R, Mendez D, Holton JM, Moriarty NW, Adams PD, Bergmann U, Sauter NK, Dobbek H, Messinger J, Zouni A, Kern J, Yachandra VK, Yano J (2021) Structural dynamics in the water and proton channels of photosystem II during the S2 to S3 transition. Nat Commun 12(1):6531. https://doi.org/10.1038/s41467-021-26781-z

Ishikita H, Knapp EW (2004) Variation of Ser-L223 hydrogen bonding with the Q_B redox state in reaction centers from *Rhodobacter sphaeroides*. J Am Chem Soc 126(25):8059–8064. https://doi.org/10.1021/ja038092q

Ishikita H, Saito K (2014) Proton transfer reactions and hydrogen-bond networks in protein environments. J R Soc Interface 11(91):20130518. https://doi.org/10.1098/rsif.2013.0518

Itoh SG, Damjanović A, Brooks BR (2011) pH replica-exchange method based on discrete protonation states. Proteins: Struct Funct Bioinform 79(12):3420–3436. https://doi.org/10.1002/prot.23176

Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 29(11):1859–1865. https://doi.org/10.1002/jcc.20945

Kaila VRI (2018) Long-range proton-coupled electron transfer in biological energy conversion: towards mechanistic understanding of respiratory complex I. J R Soc Interface 15:141. https://doi.org/10.1098/rsif.2017.0916

Kaila VR, Verkhovsky MI, Wikstrom M (2010) Proton-coupled electron transfer in cytochrome oxidase. Chem Rev 110(12):7062–7081. https://doi.org/10.1021/cr1002003

Kaila VR, Wikstrom M, Hummer G (2014) Electrostatics, hydration, and proton transfer dynamics in the membrane domain of respiratory complex I. Proc Natl Acad Sci USA 111(19):6988–6993. https://doi.org/10.1073/pnas.1319156111

Kaur D, Cai X, Khaniya U, Zhang Y, Mao J, Mandal M, Gunner MR (2019) Tracing the pathways of waters and protons in photosystem II and cytochrome c oxidase. Inorganics 7:14. https://doi.org/10.3390/inorganics7020014

Kaur D, Khaniya U, Zhang Y, Gunner MR (2021a) Protein motifs for proton transfers that build the transmembrane proton gradient. Front Chem 9:660954. https://doi.org/10.3389/fchem.2021.660954

Kaur D, Zhang Y, Reiss KM, Mandal M, Brudvig GW, Batista VS, Gunner MR (2021b) Proton exit pathways surrounding the oxygen evolving complex of photosystem II. Biochim Biophys Acta Bioenerg 1862:148446

Khaniya U, Gupta C, Cai X, Mao J, Kaur D, Zhang Y, Singharoy A, Gunner MR (2020) Hydrogen bond network analysis reveals the pathway for the proton transfer in the E-channel of *T. thermophilus* Complex I. Biochim Biophys Acta Bioenerg 1861(10):148240. https://doi.org/10.1016/j.bbabio.2020.148240

Khaniya U, Mao J, Wei RJ, Gunner MR (2022) Characterizing protein protonation microstates using monte carlo sampling. J Phys Chem B. https://doi.org/10.1021/acs.jpcb.2c00139

Kim J, Mao J, Gunner MR (2005) Are acidic and basic groups in buried proteins predicted to be ionized? J Mol Biol 348:1283–1298

Klingen AR, Ullmann GM (2004) Negatively charged residues and hydrogen bonds tune the ligand histidine pK_a values of Rieske iron-sulfur proteins. Biochemistry 43:12383–12389

Krammer EM, Till MS, Sebban P, Ullmann GM (2009) Proton-transfer pathways in photosynthetic reaction centers analyzed by profile hidden Markov models and network calculations. J Mol Biol 388(3):631–643. https://doi.org/10.1016/j.jmb.2009.03.020

LeBard DN, Matyushov DV (2008) Glassy protein dynamics and gigantic solvent reorganization energy of plastocyanin. J Phys Chem B 112(16):5218–5227. https://doi.org/10.1021/jp709586e

Liguori N, Campos SRR, Baptista AM, Croce R (2019) Molecular anatomy of plant photoprotective switches: the sensitivity of PsbS to the environment, residue by residue. J Phys Chem Lett 10(8):1737–1742. https://doi.org/10.1021/acs.jpclett.9b00437

Lu J, Gunner MR (2014) Characterizing the proton loading site in cytochrome c oxidase. Proc Natl Acad Sci USA. https://doi.org/10.1073/pnas.1407187111

Mao J, Hauser K, Gunner MR (2003) How cytochromes with different folds control heme redox potentials. Biochemistry 42(33):9829–9840. https://doi.org/10.1021/bi027288k

Mayne CG, Muller M, Tajkhorshid E (2015) Parameterizing small molecules using the force field toolkit (fftk). University of Illinois at Urbana-Champaign, Illinois

Mercado R, Bjerrum EJ, Engkvist O (2022) Exploring graph traversal algorithms in graph-based molecular generation. J Chem Inf Model 62(9):2093–2100. https://doi.org/10.1021/acs.jcim.1c00777

Merchant S, Sawaya MR (2005) The light reactions: a guide to recent acquisitions for the picture gallery. Plant Cell 17(3):648–663. https://doi.org/10.1105/tpc.105.030676

Okamura MY, Paddock ML, Graige MS, Feher G (2000) Proton and electron transfer in bacterial reaction centers. Biochim Biophys Acta 1458:148–163. https://doi.org/10.1016/s0005-2728(00)00065-7

Paddock ML, Graige MS, Feher G, Okamura MY (1999) Identification of the proton pathway in bacterial reactions centers: inhibition of proton transfer by binding of Zn^{2+} or Cd^{2+}. Proc Natl Acad Sci USA 96:6183–6188. https://doi.org/10.1073/pnas.96.11.6183

Pahari S, Sun L, Alexov E (2019) PKAD: a database of experimentally measured pKa values of ionizable groups in proteins. Database (Oxford). https://doi.org/10.1093/database/baz024

Rabenstein B, Ullmann GM, Knapp EW (1998) Calculation of protonation patterns in proteins with structural relaxation and molecular ensembles-application to the photosynthetic reaction center. Eur Biophys J 27:626–637

Reedy CJ, Gibney BR (2004) Heme protein assemblies. Chem Rev 104:617–649

Sarkar A, Gupta PL, Roitberg AE (2019) pH-Dependent conformational changes due to ionizable residues in a hydrophobic protein interior: the study of L25K and L125K variants of SNase. J Phys Chem B 123(27):5742–5754. https://doi.org/10.1021/acs.jpcb.9b03816

Sazanov LA (2015) A giant molecular proton pump: structure and mechanism of respiratory complex I. Nat Rev Mol Cell Biol 16(6):375–388. https://doi.org/10.1038/nrm3997

Sebban P, Maroti P, Schiffer M, Hanson DK (1995) Electrostatic dominoes: long distance propagation of mutational effects in photosynthetic reaction centers of *Rhodobacter capsulatus*. Biochemistry 34:8390–8397

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11):2498–2504. https://doi.org/10.1101/gr.1239303

Sharpe MA, Ferguson-Miller S (2008) A chemically explicit model for the mechanism of proton pumping in heme-copper oxidases. J Bioenerg Biomembr 40(5):541–549. https://doi.org/10.1007/s10863-008-9182-6

Song YF, Mao JJ, Gunner MR (2009) MCCE2: improving protein pK_a calculations with extensive side chain rotamer sampling. J Comput Chem 30(14):2231–2247. https://doi.org/10.1002/jcc.21222

Stowell MHB, McPhillips TM, Rees DC, Soltis SM, Abresch E, Feher G (1997) Light-induced structural changes in photosynthetic reaction center: implications for mechanism of electron-proton transfer. Science 276:812–816

Teixeira VH, Soares CM, Baptista AM (2002) Studies of the reduction and protonation behavior of tetraheme cytochromes using atomic detail. J Biol Inorg Chem 7(1):200–216. https://doi.org/10.1007/s007750100287

Torabifard H, Panahi A, Brooks CL 3rd (2020) M2 amphipathic helices facilitate pH-dependent conformational transition in influenza A virus. Proc Natl Acad Sci USA 117(7):3583–3591. https://doi.org/10.1073/pnas.1913385117

Ullmann GM, Knapp EW (1999) Electrostatic models for computing protonation and redox equilibria in proteins. Eur Biophys J 28(7):533–551. https://doi.org/10.1007/s002490050236

Ullmann GM, Noodleman L, Case DA (2002) Density functional calculation of pK_a values and redox potentials in the bovine Rieske iron-sulfur protein. J Biol Inorg Chem 7:632–639

Vassiliev S, Zaraiskaya T, Bruce D (2012) Exploring the energetics of water permeation in photosystem II by multiple steered molecular dynamics simulations. Biochim Biophys Acta 1817(9):1671–1678. https://doi.org/10.1016/j.bbabio.2012.05.016

Vila-Vicosa D, Reis P, Baptista AM, Oostenbrink C, Machuqueiro M (2019) A pH replica exchange scheme in the stochastic titration constant-pH MD method. J Chem Theory Comput 15(5):3108–3116. https://doi.org/10.1021/acs.jctc.9b00030

Vilhjalmsdottir J, Albertsson I, Blomberg MRA, Adelroth P, Brzezinski P (2020) Proton transfer in uncoupled variants of cytochrome c oxidase. FEBS Lett 594(5):813–822. https://doi.org/10.1002/1873-3468.13679

Vinyard DJ, Brudvig GW (2017) Progress toward a molecular mechanism of water oxidation in photosystem II. Annu Rev Phys Chem 68:101–116. https://doi.org/10.1146/annurev-physchem-052516-044820

Wei RJ, Zhang Y, Mao J, Kaur D, Khaniya U, Gunner MR (2022) Comparison of proton transfer paths to the QA and QB sites of the *Rb. sphaeroides* photosynthetic reaction centers. Photosynth Res. https://doi.org/10.1007/s11120-022-00906-x

Wraight CA (2004) Proton and electron transfer in the acceptor quinone complex of photosynthetic reaction centers from *Rhodobacter sphaeroides*. Front Biosci 9:309–337

Zhang Y, Haider K, Kaur D, Ngo VA, Cai X, Mao J, Khaniya U, Zhu X, Noskov S, Lazaridis T, Gunner MR (2021) Characterizing the water wire in the gramicidin channel found by monte carlo sampling using continuum electrostatics and in molecular dynamics trajectories with conventional or polarizable force fields. J Comput Biophys Chem 20(02):111–130. https://doi.org/10.1142/s2737416520420016

Zheng Y, Cui Q (2017) Microscopic mechanisms that govern the titration response and pKa values of buried residues in staphylococcal nuclease mutants. Proteins 85(2):268–281. https://doi.org/10.1002/prot.25213

Zheng Z, Gunner MR (2009) Analysis of the electrochemistry of hemes with E_ms spanning 800 mV. Proteins 75:719–734. https://doi.org/10.1002/prot.22282