Identifying the Proton Loading Site cluster in the Ba3 Cytochrome c Oxidase that loads and traps protons

Xiuhong Cai[1,2], Chang Yun Son[3,4], Junjun Mao[1], Divya Kaur[1,5], Yingying Zhang[1,2], Umesh Khaniya[1,2], Qiang Cui[6], M. R. Gunner[1,2,*]

1. Department of Physics, City College of New York, 160 Convent Avenue New York, NY, 10031, USA
2. Department of Physics, Graduate Center, City University of New York, 365 Fifth Avenue, New York, NY 10016, USA
3. Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, 91125, USA
4. Current Address: Department of Chemistry, Pohang University of Science and Technology (POSTECH), Pohang, 37673, Korea
5. Department of Chemistry, Graduate Center, City University of New York, 365 Fifth Avenue, New York, NY 10016, USA
6. Department of Chemistry & Department of Biomedical Engineering & Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA, 02215, USA

* To whom correspondence should be addressed: mgunner@ccny.cuny.edu.

Abstract

Cytochrome c Oxidase (CcO) is the terminal electron acceptor in aerobic respiratory chain, reducing O_2 to water. The released free energy is stored by pumping protons through the protein, maintaining the transmembrane electrochemical gradient. Protons are held transiently in a proton loading site (PLS) that binds and releases protons driven by the electron transfer reaction cycle. Multi-Conformation Continuum Electrostatics (MCCE) was applied to crystal structures and Molecular Dynamics snapshots of the B-type *Thermus Thermophilus* CcO. Six residues are identified as the PLS, binding and releasing protons as the charges on heme *b* and the binuclear center are changed: the heme a_3 propionic acids, Asp287, Asp372, His376 and Glu126B. The unloaded state has one proton and the loaded state two protons on these six residues. Different input structures, modifying the PLS conformation, show different proton distributions and result in different proton pumping behaviors. One loaded and one unloaded protonation states have the loaded/unloaded states close in energy so the PLS binds and releases a proton through the reaction cycle. The alternative proton distributions have state energies too far apart to be shifted by the electron transfers so are locked in loaded or unloaded states. Here the protein can use active states to load and unload protons, but has nearby trapped states, which stabilize PLS protonation state, providing new ideas about the CcO proton pumping mechanism. The distance between the PLS residues Asp287 and His376 correlates with the energy difference between loaded and unloaded states.

Keywords: Monte Carlo Sampling; Cytochrome c Oxidase; proton affinity; MCCE

Abbreviations: CcO, Cytochrome c Oxidase**;** BNC, Binuclear Center; PLS, proton loading site**;** MD, molecular dynamics; MCCE, Multi-Conformation Continuum Electrostatics; PRA_{a3}: propionic acid A of heme a_3; PRD_{a3}, propionic acid D of heme a_3; PRA_b, propionic acid A of heme *b*; PRD_b, propionic acid D of heme *b*; MS, microstate.

1. Introduction

Cytochrome c oxidase (CcO), a member of the heme-copper oxidase superfamily, is a proton pump, driven by the chemical energy produced by reduction of oxygen to water [1].The proton is pumped through chains of waters and amino acids from the N-side (high pH) to the P-side (low pH). Pumping protons uphill is a complex process, requiring changes in proton affinity and proton connections within the working protein.

CcO, Complex I [2] and bacteriorhodopsin [3] are representative proton pumping proteins. Proton pumps move protons through the entire membrane embedded protein driven by energy of a photon or a redox reaction [4,5]. Pumps have proton loading sites (PLS), a residue or group of residues whose proton affinity changes dramatically through the reaction cycle, so a proton is loaded, held and then released through the reaction cycle. A requirement for a pump to function is that the PLS proton affinity must change so that its effective pK_a shifts substantially, from several pH units above the pH for proton loading to several pH units below the pH for proton unloading [6,7]. Smaller changes in affinity or those where the change of effective pK_as in different states that do not cross the pH of the surroundings will not cause the PLS to alter protonation state. The electrochemical gradient across the membrane, favors wrong way proton transfer from P- to N-side, so for efficient pumping, when protons are loaded the proton pathway must connect to the N-side and opened only to the P-side when protons are released [8]. Thus, the study of proton pumps must consider the energetics of proton binding sites as well as the connectivity of proton transfer pathways.

In CcO the energy required is provided by the reduction of oxygen at the Binuclear center (BNC) located in the center of the protein. The BNC consists of a heme group and a copper site and a covalently-linked Tyr residue [7,9,10]. In the overall redox reaction cycle CcO reduces one oxygen molecule to water, driving a reaction cycle that results in uptake of 4 protons from the N-side for chemistry and additional protons pumped from N- to P-side [7,9]:

$$O_2 + 4Cyt\ c_P^{2+} + 4(1+\eta)H_N^+ \rightarrow 2H_2O + 4Cyt\ c_P^{3+} + 4\eta H_P^+, \quad (1)$$

where η is the number of protons pumped from the N- to P-side of the membrane/electron [11].

CcO is divided into three families denoted by A, B and C [11–13]. The A-type CcO is present in eukaryotes and bacteria. B- and C-type CcO are found in bacteria and archaea that live at lower oxygen levels [12]. The different families differ in the nature of the active site hemes (heme *a*, *b*, or *c*), the number of N-side proton pumping pathways (one or two) and the number of protons pumped/electron (η).

The A-type *aa3* CcO is the best studied, with Cu_A as the electron acceptor from cytochrome c, a heme *a* and heme a_3 and Cu_B as the catalytic Binuclear Center (BNC) [5,10,14–16]. The *ba3* CcO is a B-type CcO. It uses heme *b* instead of heme *a* to accept electrons from Cu_A on the P-side and transfer them to the BNC. The isolated heme *b* has an E_m of-120mV while heme *a* -20mV [17], which can increase the driving force for electron transfer into the BNC. The number of protons/electron, η, is ≈1 for A-type CcO [15,18–20] while it is ≈0.5 for the B and C type CcO [21–23].

In every family of CcO the reduction of O_2 takes four sequential reductions of the BNC, denoted the redox states F, O, E, R [1,24–26]. In the fully reduced R state (Fe_{a3}^{II}, Cu_B^I, Tyr-OH), O_2 is bound to the iron atom of heme a_3. It is at this stage that 4 electron reduction of O_2 occurs in the BNC, leading to the most oxidized protein (F). Then one electron and one chemical proton is sequentially transferred to the BNC in each redox transition from O to E and back to R. (Figure SI.1, Table SI.1).

The PLS has been studied in *Rb. sphaeroides* [25,27], *Paracoccus denitrificans* [28–30], bovine[31–33] and yeast A-type CcO [34,35]. Various studies suggested the propionic acids of heme a_3 or the Histidine ligand to Cu_B to be the potential loading sites for A-type *aa3* CcO [28,29,32,36,37]. Calculations using methods similar to the ones that will be applied here proposed that the proton loading site is a cluster of residues instead of a single site, with the two heme propionic acids of heme a_3 playing the major role [25]. An analysis of the hydrogen bond network of A-type CcO also identified a cluster of residues that include, but extend beyond, the proton holding PLS residues on the P-side that can lead to proton egress [8].

As in the A-type CcO [8] the *ba3* type CcO has an extended group of buried charged and polar residues on the P-side of the BNC. In the *ba3* CcO, the propionate A of heme a_3 and nearby residues have been suggested to be the PLS [38–43]. Mutations of Asp372, His376, Asp287 disrupted or reduced proton pumping [40]. A conserved water molecule between propionates A and D of heme a_3 [39,40,44] or a water-dimer near the Cu_B site in the BNC [45] have also been proposed to be part of the proton pumping pathway. One model [46] highlights a possible role of His376, suggesting that its rotation moves a proton from the propionate A of heme a_3 or Asp372 onto Glu126B and onward to the P-side surface. Comparison of the pumping mechanisms of the members of the CcO family can help us understand the constraints on the transiently proton trapping element of the proton pumping cycle.

In A-type CcO, four protons are pumped with the four electrons needed for one O_2 reduction. After much discussion of alternative patterns, one proton is believed to be pumped in each redox reaction [28,47–54]. To pump one proton each time the BNC is reduced, the proton affinity of the PLS must change four times during the redox cycle from being above to below the solution pH. A Monte Carlo analysis of proton binding in *Rb. sphaeroides* A-type CcO through the reaction cycle showed protons could be loaded and unloaded in each redox state [25].

Here, CcO structures are obtained from crystal structures from the Protein Data Bank and snapshots from Molecular Dynamics (MD) trajectories conducted in different redox states with different protonation states of possible PLS residues. Continuum electrostatics based Monte Carlo analysis (Multi-Conformation Continuum Electrostatics, MCCE) is applied, allowing the protonation states of the whole protein to come to equilibrium with a series of imposed redox states that models the reaction cycle in 136 different starting structures. This approach allows us to probe the coupling between the redox cycle and the protonation states of the protein in diverse structures. The method does not pre-select the PLS residues.

A cluster of residues is identified whose average protonation increases when the BNC is reduced and diminished when a substrate proton is added to the BNC. The *ba3* CcO PLS is in a location similar to that found in the *aa3* CcO [25]. However, the propionic acids are the key loading sites in the *aa3* protein, while here we find a more delocalized PLS made of the propionic acids of heme a_3, Asp287, Asp372, His376 and Glu126B (Figure 1). The unloaded PLS cluster has one proton and the loaded state has two. The changes in heme *b* redox state and the redox and protonation states of the BNC always tune the proton affinity of the PLS cluster by 5-6 kcal/mol. However, the free energy difference between the protonated and deprotonated PLS states, which corresponds to the proton affinity of the cluster, can be too large for the reaction cycle to change the protonation state. Thus, individual structures are found locked in loaded (2 protons) or unloaded (1 proton) states, or in ones that can actively bind and release protons. Locked unloaded and active pumping states were also seen in our earlier study of *aa3* CcO, but at the time we could not identify the underlying cause [25]. The distribution of protons and the hydrogen bonding pattern amongst the residues in the PLS cluster are different in the

active and locked structures. Thus, CcO may move from an active configuration to bind or release protons to a locked configuration to preserve the loaded or unloaded state. Compared with A type CcO, B type CcO only has half of the yield of pumped protons, with only ≈2 pumped/O_2. No input structure binds more than 3.5 protons into the PLS through the whole redox cycle and many bind substantially less.

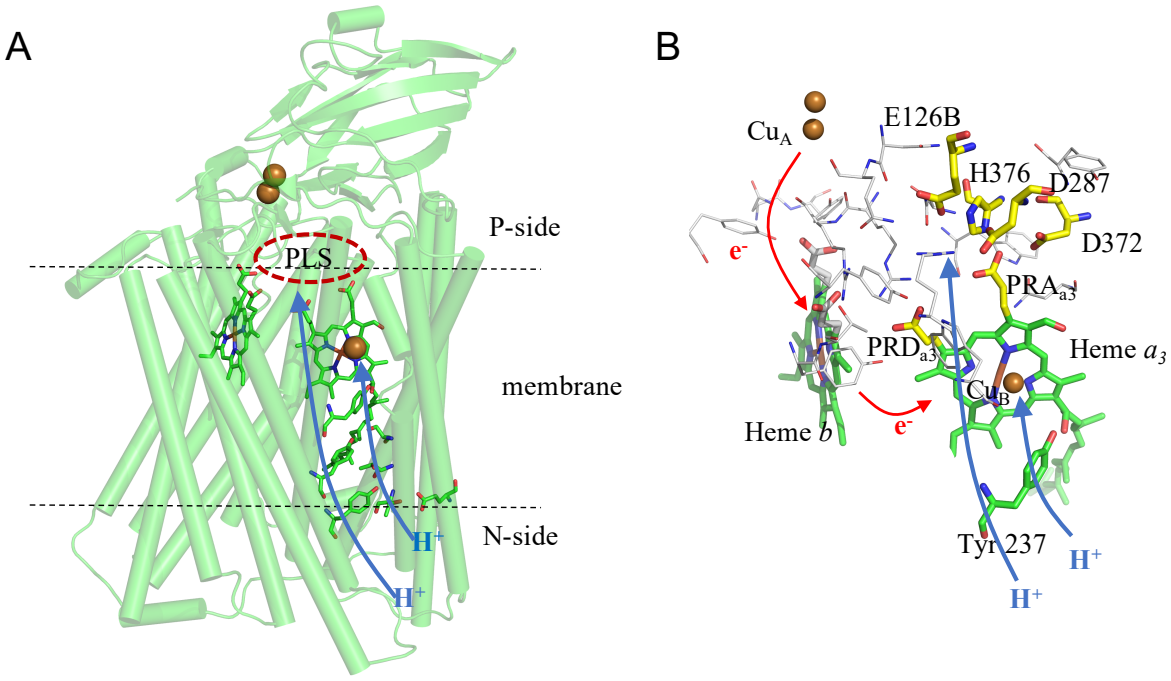

Figure 1 A) B type CcO from Thermus Thermophilus structure PDB:3S8F. The hemes are shown with sticks, heme b on the left and heme a_3 (with a long tail) on the right. The di-copper Cu_A and Cu_B in the BNC are shown as gold spheres. The K channel [39,55] residues are shown as sticks to the bottom right. The region where the PLS is expected to be located is circled. B) Electrons are transferred from a P-surface bound cytochrome c (not shown) to Cu_A then to heme b which reduces the BNC (heme a_3 and Cu_B). Protons needed for chemistry are added to the oxygen ligands bound on heme a_3 and Cu_B as a product of O_2 reduction. The six labeled residues that are found to function as the PLS are shown as sticks with yellow carbons. Residues that are connected by hydrogen bonds to these residues shown as grey lines while the propionates of heme b are shown as grey sticks.

2. Methods and Materials

2.1 MD simulation

MD simulations are prepared at seven different redox/protonation states, starting from the crystal structure, 3S8F, of *Thermus Thermophilus* CcO [56] (Table 1). The states were initially chosen given the proposed reaction cycle where the P to F transition is coupled to proton pumped from N-side to the PLS [9]. Therefore, CcO was simulated with no protons added to the PLS region and with protons on the Propionic acid A of heme a_3 or on Asp372. In addition, CcO with protons, on Asp372 and His376, were prepared, as MCCE analysis showed these to be the preferred proton location in simulations starting with the crystal structures. The MD simulation protocols are identical to those described previously [57].

Table 1 Redox and protonation states prepared for MD simulation

Trajectory	State	Designation	Details	Color[1]
1	F_eMS0	P^3	PRA_{a3}^- , Cu_B(II)-OH^-	Orange
2	F_eMS2	$P^{3'}$	$PRA_{a3}H$, Cu_B(II)-OH^-	Purple
3	F_{+e}MS0	F^3	PRA_{a3}^-, Cu_B(II)-H_2O	Cyan
4	F_{+e}MS2	$F^{3'}$	$PRA_{a3}H$, Cu_B(II)-H_2O	Green
5	F_{+e}MS5	$F^{3''}$	D372H, Cu_B(II)-H_2O	Yellow
6	F_{+e}MS14	$F^{3'''*}$	$PRA_{a3}H$, D372H, Cu_B(II)-H_2O	Black
7	F_{+e}MS16	$F^{3'''}$	H376H, D372H, Cu_B(II)-H_2O	Red
8			Crystal structures	Blue

All trajectories have Ferryl (Fe (IV)) heme a_3 with an O^{2-} ligand, with Heme *b* reduced and Tyr237 ionized (YO^-). State: identifies the trajectories with the most similar MCCE substate (Figure SI.1 and Table SI.1). Designation: trajectory designation in earlier work that defined the MD running conditions [9]. P^3 has Cu_B(II)-OH^- and a reduced heme *b*, which occurs before the F_e substate in Table SI.1, thus before heme *b* is oxidized, while F^3 has a proton added to the ligand on Cu_B. All trajectories are 50 ns long. MS 0 has no proton on any of the six PLS residues (charge -5); MS 2 has a single proton on PRA_{a3} while MS 5 has a single proton on Asp372; MS 14 PLS has two protons, one on PRA_{a3} and Asp372; in MS 16 His376 and Asp372 are protonated (Table SI.2). Color: connects the results presented in Figure 2 and 3 with the trajectory that is the source of the snapshots that initiate the MCCE calculations.

2.2 MCCE simulation

Twelve crystal structures of the *ba3* CcO from *Thermus Thermophilus* (PDB codes: 1EHK, 1XME, 3EH3, 3EH4, 3EH5, 3S3A, 3S3B, 3S3C, 3S3D, 3S8F, 4GP4, 4GP5) [38,56,58–61] from the Protein Data Bank and MD snapshots at different redox/protonation states (Table 1) are analyzed by MCCE simulation. In the beginning, MD snapshots are selected by their time point in the trajectory. To guarantee the diversity of input structures, the ≈3,500 snapshots from the seven MD trajectories were then clustered by the coordinates of the PLS cluster (propionic acids of heme a_3, Asp287, Asp372, His376, Glu126B) using MD Analysis [62–64] generating 124 clusters. Each of the trajectories generates 13-24 clusters with from 1 to79 snapshots/cluster. One snapshot representing the centroid of each cluster is chosen for MCCE analysis of the proton pumping behavior.

MCCE [65] can generate multiple conformers for residue side chains as well as the protonation and redox state for each residue and cofactor. Monte Carlo sampling allows the entire system to come to equilibrium. Here conformational sampling was restricted to isosteric side chain conformations with only the cofactor redox and protonation states imposed for each

state in the reaction cycle (Table SI.1 and Figure SI.1). These calculations keep the heavy atoms determined by the input MD snapshot or crystal structure but samples His tautomers, hydroxyl proton positions and Asn and Gln termini positions as well as the protonation state for each acidic and basic residue. Explicit water molecules are removed and replaced with implicit solvent [66]. The hydrophobic portion of the lipid bilayer from the MD simulation is retained to serve as a low dielectric surroundings in MCCE simulations (Figure SI.2A). The membrane was added to the crystal structure coordinates by alignment to an equilibrium MD snapshot with PyMol. The atomic charges for Cu_A, Cu_B, heme a_3 are given in the earlier publications [25,67,68]. Atomic charges for heme *b* are benchmarked in MCCE [68] (Table SI.3). Approximately 4.5 million steps of MC sampling are averaged to determine the equilibrated protonation state for each structure in each defined reaction state.

2.2.1 Modeling the changes in CcO redox and protonation state in MCCE

Independent MCCE calculations are carried out for 136 MD snapshots or crystal structures in twelve redox and protonation states for the BNC, Cu_A and heme *b* (Table SI.1 and Figure SI.1). The four redox states are designated R-F-O-E [9]. Each redox state (X, where X can be R, F, O or E) starts with electron transfer to the BNC (forming the state designated X_e), followed by proton addition to the BNC (X_p) and then electron transfer to heme *b* (X_{+e}). Protons are added to the OH^- or $O^=$ bound to Heme a_3 or Cu_B or to Tyr237. The BNC states are: R: (Fe_{a3}^{II}, Cu_B^I, Tyr-OH); F:(Fe_{a3}^{IV}=$O^=$, Cu_B^{II}-H_2O, Tyr-O^-); O:(Fe_{a3}^{III}-OH^-, Cu_B^{II}-OH^-, Tyr-OH); E: (Fe_{a3}^{III}-OH^-, Cu_B^I, Tyr-OH). Heme *b* is reduced in X_{+e} and oxidized in X_e or X_p. The total change on heme *b* and the BNC is -1 in X_{+e} and X_e and 0 in X_p.

The Boltzmann distributions of protonation states are calculated for each MCCE redox/protonation substate in each structure. Six residues will be identified as constituting the PLS cluster (PRA_{a3}, PRD_{a3}, Asp287, Asp372, His376 and Glu126B). The PLS charge for a structure in a given CcO reaction state is the sum of the Boltzmann averaged protonation of these six residues. The average ionization state for this structure is the average PLS charge over the 12 states in the reaction cycle. The integrated proton loading into the PLS region for one reaction cycle (ΔH^+) is:

$$\Delta H^+ = \sum_{redox\ state} H^+_{max} - H^+_{min}, \quad (2)$$

where H^+_{max} is the maximum number of protons loaded into the PLS for each of the four redox states (R, F, O and E) and H^+_{min} is the minimum protonation in the previous redox state. Analyses of protons bound are obtained from MC sampling of the entire protein.

2.2.2 Energy map and proton binding energy of PLS

MC analysis samples low energy states but gives little information about states at very high energy. Thus, if the protein is found in only one protonation state, we only know that other states are at low probability. Here ≈4.5 million microstates are evaluated, which sets the limit on the likelihood of finding rare states. A more direct analysis of the microstate energy uses the enumerated energy of each protonation microstate of the six PLS residues (PRA_{a3}, PRD_{a3}, Asp287, Asp372, His376 and Glu126B) within a fixed, averaged background from the rest of the protein. The microstate energy here does not count for the entropy generating by the conformers created in MCCE while the full analysis MC analysis of the entire protein, used to analyze the proton pumping behavior above, does.

All residues in the protein except those in the PLS cluster are fixed in the MCCE Boltzmann distribution equilibrated in a given state of heme *b* and the BNC. Thus, the non-PLS

residues act on the PLS by their averaged charge and position (a mean-field approach). The interaction between PLS conformer i and the fixed, non-PLS conformers (j) in the environment, ΔG_{ij}^{mfe}, is:

$$\Delta G_{ij}^{mfe} = P_j \Delta G_{ij}, \quad (3)$$

where P_j is the Boltzmann averaged occupancy of conformer j previously determined in a free MCCE calculation for that structure in a specific reaction substate, and ΔG_{ij} is interaction between conformer i of PLS residues and conformer j of non-PLS residues.

The total energy of PLS microstate (MS), where x is the conformer combination of the six-residue PLS, is:

$$\begin{aligned}\Delta G_{MS} &= \textstyle\sum_{i\in x}\sum_j \Delta G_{ij}^{mfe} \\ &+ \textstyle\sum_{i\in x}(\Delta G_i^{bkb} + \Delta G_i^{sol}) \\ &+0.5 * \textstyle\sum_{i\in x, k\in x, i\neq k} \Delta G_{ik}, \end{aligned} \quad (4)$$

where ΔG_i^{bkb} is the interaction between the fixed protein backbone dipoles and PLS conformer i and ΔG_i^{sol} is the solvation energy of conformer i. ΔG_{ij}^{mfe} is given by Equation 3. ΔG_{ij} is the interaction between conformer i of one PLS residue and conformer k of the other five PLS residues. The factor of 0.5 corrects for the interaction between PLS residues being counted twice in the summation, See Figure SI.3 for a more complete description. The free energy of all PLS conformation and protonation microstates are determined in each structure using the ΔG_{MS} calculated in each step of the reaction cycle. We will use the minimum energy conformational microstate with the desired proton distribution to designate the energy of each of the 64 individual PLS protonation microstates. Thus, this is the energy rather than the free energy.

The PLS generally shows a loaded state with two protons loaded (a -3 net charge) and an unloaded state with one proton (a -4 net charge). The free energy difference of PLS between loaded and unloaded states in a specific structure in a specific MCCE substate is:

$$\Delta\Delta G_{UL-L} = min\Delta G_{UL} - min\Delta G_L, \quad (5)$$

where $min\Delta G_{UL}$ is the minimum energy of all microstates with a charge of -4 and the $min\Delta G_L$ is the minimum energy of all microstates with a -3 charge. A positive energy difference indicates that PLS prefers to be loaded.

3. Results

3.1 Identification of the PLS residues and characterization of the proton loading/unloading cycle

3.1.1 The highly interconnected residues that are candidates for the PLS of *ba3* CcO

The Proton Loading Site (PLS) of B-type CcO is defined as a group of residues that bind and release protons during the proton pumping cycle. Our search for the PLS will consider residues that are located between heme *b* and the BNC, near the middle of the protein, and the P-side surface (Figure 1). A functional PLS must be connected to proton pathways with alternating access to the N-side of the protein for proton loading and P-side for release. Here, the MCCE hydrogen bond analysis previously used to study the *aa3* CcO [8] finds twenty amino acid residues and the four propionic acids residues that are in the right location and are in an interconnected, water mediated, hydrogen bond network that can facilitate proton transfer through a Grotthuss mechanism in the *ba3* CcO (Figure SI.4). These PLS candidates include 4 basic amino acids (Arg225, 449, 450 and His376), 3 acidic amino acids (Asp 287, 372, and Glu126B), and the 4 propionic acids of heme *b* and heme a_3 as well as 13 polar residues (Table SI.4). The proton binding PLS residues will be identified by their changing protonation states in MCCE calculations of the proton distribution in the whole protein through the CcO redox driven proton pumping cycle [1,25,29,30,47,69–71].

3.1.2 The imposed redox state changes to model the reaction cycle

The redox state of heme *b* and the BNC redox and protonation states are altered in MCCE to represent the CcO reaction cycle. The R, F, O, E BNC redox states are each divided by a sequence of three changes (Table SI.1 and Figure SI.1). A new redox state is initiated when an electron is transferred from heme *b* to the BNC, reducing heme a_3 or Cu_B (ET). This forms state X_e (where X is R, F, O or E). Then a proton is added from the outside to one of the two product waters in the BNC (PT) to form the states designate X_p. Lastly, an electron is added to heme *b* (+E) to form X_{+e}. The electron transfer from an external *cyt.* c to Cu_A that precedes heme *b* reduction is not modeled here.

MCCE brings the protonation states and the polar proton positions of all other residues into equilibrium with the 12 redox/protonation states that define the imposed changes at heme *b* and the BNC through the reaction cycle. The protein backbone and the side chain heavy atoms are not allowed to move. Rather, structural diversity is added by analyzing 136 structures derived from crystal structures and selected MD snapshots from trajectories run in different redox and protonation states (Table 1).

3.1.3 PLS residues that change protonation state as the redox state changes

The residues monitored as the PLS candidates do indeed bind and release at most one proton when the redox and protonation states of the principle cofactors change. The two propionic acids of heme a_3, Asp287, Asp372, His376, Glu126B, either change their protonation state or interact strongly with the residues undergoing protonation changes, often changing position, so they cannot be separated from the cluster (Figure 1B and Figure 2A-2C). Thus, these six residues are proposed to function as the PLS. All other residues in the PLS candidate region remain in their standard ionization state with Asp, Glu, Arg and Lys ionized, His and Tyr neutral in MC sampling through the 12-step reaction cycle in all input structures. Small

ionization state changes are found in mostly surface residues that are disconnected from the PLS proton transfer chains. These will not be discussed further but are shown in Figure SI.2B.

3.1.4 Choice of structures for analysis of proton loading

MD trajectories allow the protein to explore conformational space with fixed redox/protonation states in a limited window of time (50 ns here). The question is what differentiates these input structures to favor different PLA protonation states. A meta-analysis was carried on the seven MD trajectories, clustering the positions of the six PLS residues, producing a set of structures that capture the range of PLS conformations [62–64]. Table 2 indicates the likelihood of the different structural categories. One structure in each of the 124 conformational clusters in addition to the 12 crystal structures, are simulated through the cycle of different BNC redox/protonation states.

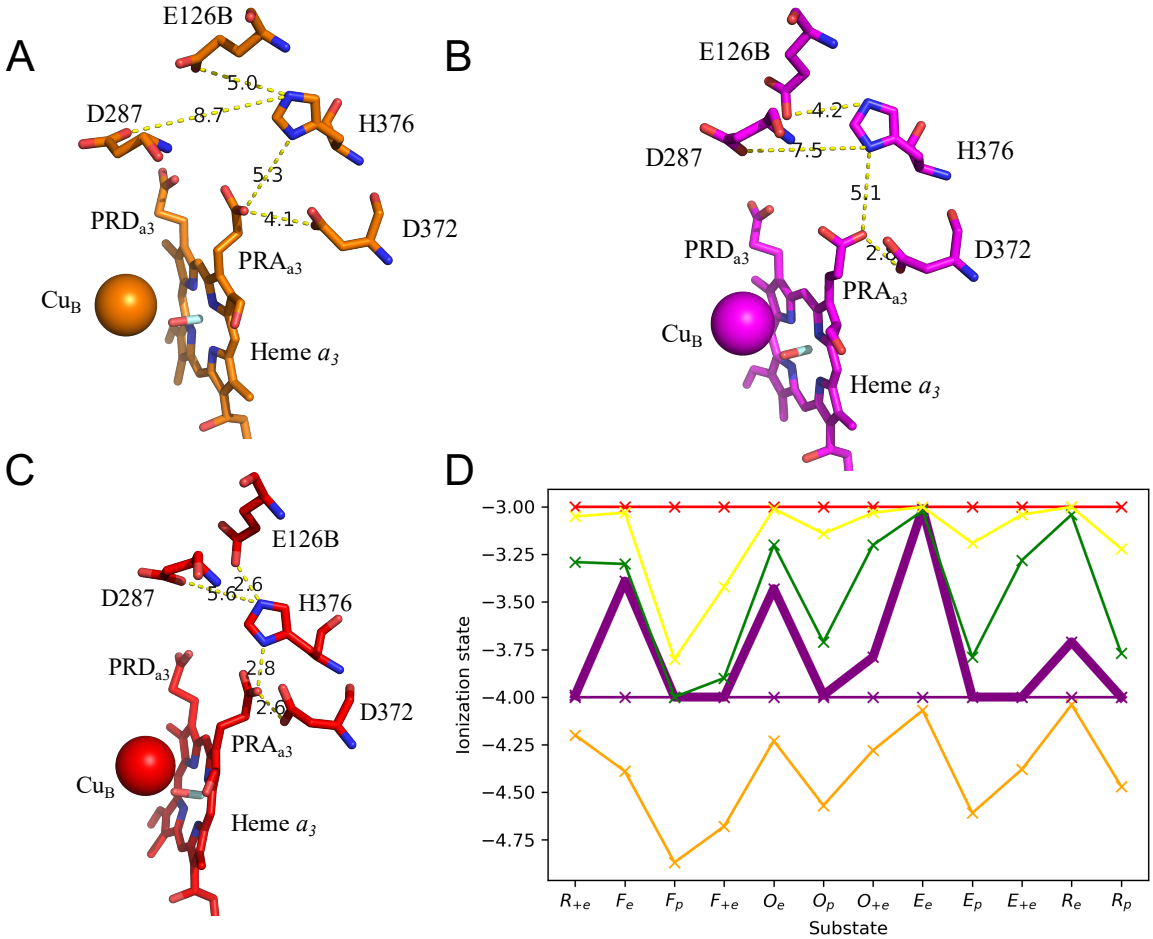

Figure 2 A-C) The position of Cu_B and Heme a_3 and the six PLS residues identified here from selected input structures showing the range of positions. The line colors identify the source of the input structure (Table 1). A) Orange, Has proton uptake between -4 and -5; B: Purple; Active proton loading with a charge between -3 and -4; C: Red, Fixed loaded with a -3 charge. D) The net charges of the six-residue PLS in equilibrium with each of the 12 MCCE substates that model the reaction cycle for different structures. The orange, thicker purple and red lines show the net charges of structure 2A, 2B and 2C respectively. The line color identifies the source of the input structure defined in Table 1. Blue: crystal structures; Red: F_{+e}MS16; Black: F_{+e}MS14; Yellow: F_{+e}MS5; Green: F_{+e}MS2, Cyan: F_{+e}MS0, Purple: F_eMS2, Orange: F_eMS0. Given that X can be the R, F, O or E redox state, X_{+e} indicates that an electron is added to heme b; X_e that the electron is transferred from heme b to heme a_3 or Cu_B in the BNC; X_p

that a proton is added to one of the product waters bound in the BNC. The imposed charge on heme b and the BNC is -1 in X_e and X_{+e} and zero in X_p substates. Heme b is reduced in X_{+e} and oxidized otherwise. The thicker purple line is described as an example in the text.

Figure 2D provides examples of how the PLS charge changes through the imposed 12 step reaction cycle, showing the results of calculations initiated with different structures In MC sampling a defined microstate gives each residue a specific protonation state and so it has an integer charge. With 5 acids and one base the charge of the PLS can vary from -5, with no protons bound, to +1, with all residues protonated. A non-integer charge reflects that accepted microstates have different protonation states. The R_{+e} substate (the BNC in the protonated R state with heme *b* reduced) is taken as an arbitrary starting state. The fractional proton binding is the Boltzmann averaged change in charge of the six residues in the PLS cluster.

The bold line in Figure 2D will be described as an example of the proton loading/unloading behavior. The six PLS residues have a net charge of -4.0 in the R_{+e} substate. Thus, one of the residues has bound a proton. When the electron is moved from heme *b* to the BNC (generating the F_e substate) MCCE calculates that on average 0.6 proton is loaded into the PLS. Proton binding to the BNC (F_p), leads to proton unloading and the recovery of the -4.0 net charge. There is little change in PLS protonation when heme *b* is reduced (F_{+e}). The cycle continues through O, E and back to the R state, with each redox state having a different distribution of electrons and protons in the BNC (Table SI.1).

Figure 2D shows examples of the range of behaviors found with different input structures. The PLS never binds more than one proton during the proton uptake phase of the reaction cycle, however it can bind less. Many structures change their protonation state in concert with the imposed charge on heme *b* and the BNC, moving between having one and two protons bound. These are considered active. Some structures always keep 2 protons bound, generating a charge of -3, independent of the imposed redox/protonation states. These structures are locked in the loaded state (Table 2). Others maintain a charge of -4, with one bound proton, which are locked in the unloaded state. There is a small subset of structures that transition from a charge of -5 to -4, which will be denoted active/empty.

The PLS charge is somewhat different in the same substate (X_e, X_p, X_{+e}) in different redox states (R,F,O,E). E_e (65%) or R_e (68%) has the most protons loaded (these are double counted when E_e and R_e have the same number of protons bound), while in 97% of the structures the F_p substate has the fewest protons bound.

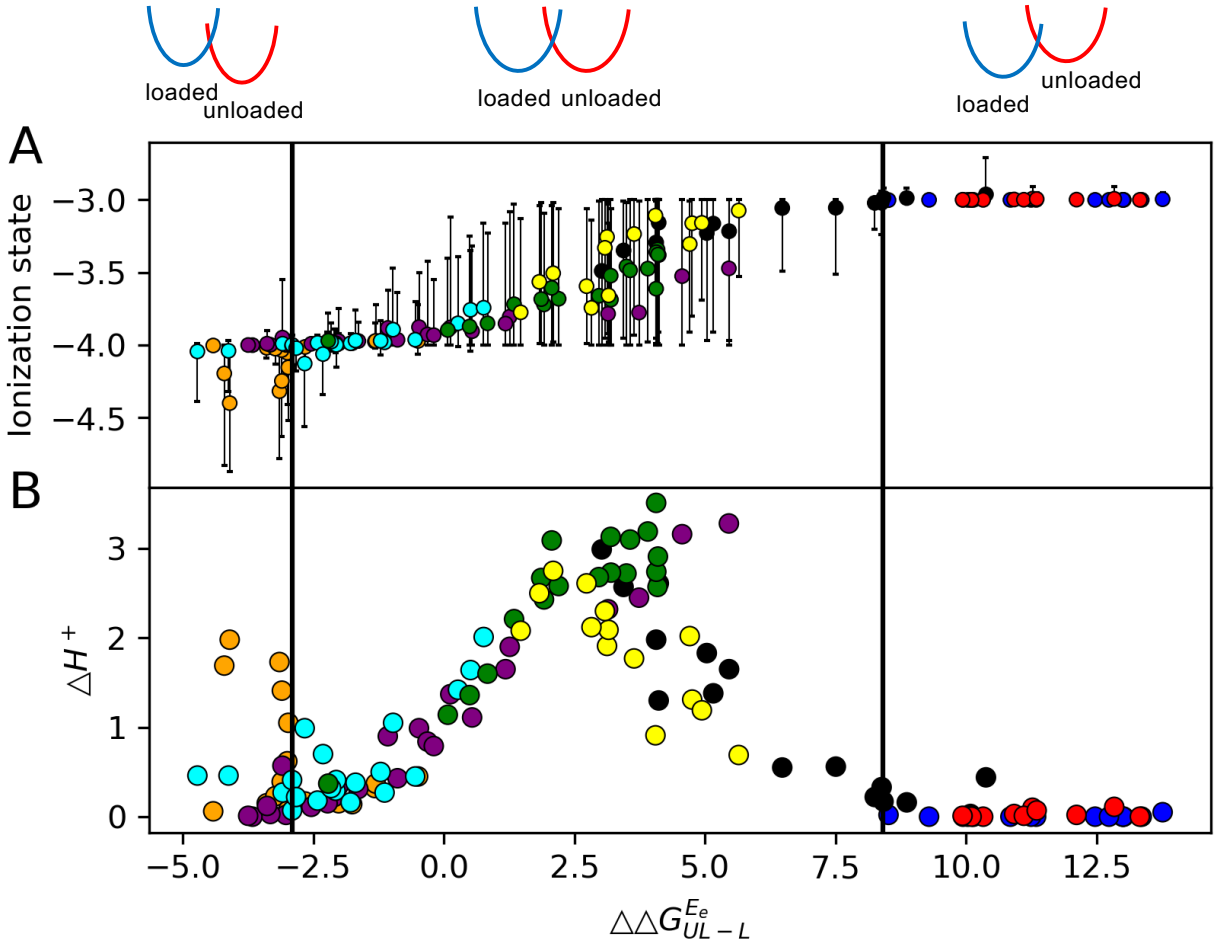

Figure 3 The proton loading behavior as a function of the energy difference between the lowest energy unloaded (-4 charge) and loaded (-3 charge) microstates in the E_e state of heme b and the BNC, $\Delta\Delta G^{Ee}_{UL-L}$ (Equation 5). More positive $\Delta\Delta G^{Ee}_{UL-L}$ favors proton loading. Protonation changes are determined from MC sampling of the whole protein. The schematic at the top shows that on the left the unloaded state is favored, in the middle the two charge states have similar energy and on the right the loaded state is at lower energy. A) Dots show the average of the MC determined ionization state in the 12 imposed states in the reaction cycle; Line: minimum to maximum charge found in the entire 12 step cycle. B) Dots are the integrated proton uptake through the full reaction cycle (Equation 2). The dot color identifies the source of the input structure defined in Table 1. Blue: crystal structures; Red: $F_{+e}MS16$; Black: $F_{+e}MS14$; Yellow: $F_{+e}MS5$; Green: $F_{+e}MS2$, Cyan: $F_{+e}MS0$, Purple: F_eMS2, Orange: F_eMS0. MS identifies the PLS protonation microstate assignments in MD. MS 0 has no protons in the PLS (charge -5); MS 2 has a single proton on PRA_{a3} while MS 5 has a single proton on Asp372; MS 14 has two protons in the PLS on PRA_{a3} and Asp372; in MS 16 they are on His376 and Asp372.

3.2 The input structure determines if the PLS can respond to the changes through the reaction cycle or be locked in loaded or unloaded state

3.2.1 Characterization of structures by their proton loading characteristics

The pumping behavior depends on how the PLS proton affinity changes through the imposed 12 step reaction cycle. However, it is also depending on the overall proton affinity of the PLS, which is a function of the input structure. Structures are categorized by the free energy difference between the loaded (L, -3 PLS charge, 2 protons bound) and unloaded (UL, -4 charge, one proton bound) states at a single redox state, $\Delta\Delta G_{UL\text{-}L}$ (Equation 5). When $\Delta\Delta G_{UL\text{-}L}$ is

positive, the PLS favors being loaded. Figures 3A and B use $\Delta\Delta G_{UL-L}^{Ee}$ for the x-axis since the E_e substate has the most protons loaded through the reaction cycle in almost all active structures. For simplicity the unique lowest energy unloaded, and loaded microstates are compared, rather than the Boltzmann averaged energy, of the microstates with that charge.

Figure 3A characterizes the total PLS charge determined by MC sampling for each structure. The bottom of the vertical line is the charge in the most unloaded state in the reaction cycle (usually F_p) while the top is the charge in the most loaded state (E_e or R_e). The dot is the averaged charge over the 12-step reaction cycle, giving the weighted average of the PLS proton affinity. A long vertical line indicates protons are bound and released. No line shows that the PLS is locked in the loaded or unloaded state. For the structure followed by the bold line in Figure 2D the line extends between -4 and -3.

ΔH^+, the integrated proton uptake, shows the ability of the system to respond to the imposed charge change through our reaction cycle (Figure 3B). If there is no back-proton transfer this is the maximum number of protons pumped through the reaction cycle for the given structure. The integrated proton uptake of the PLS cluster for the four-electron reduction cycle is the sum of the protons loaded in each of the four redox states, adding the difference in protonation between X_e and X_p for each of the four BNC redox states (Equation 2). Structures have zero proton uptake if they are fixed loaded or unloaded. For the structure followed by the bold line in Figure 2D the integrated proton uptake is 2.5 protons. The maximum ΔH^+ here is 3.5 (Table 2). It is never 4.

3.2.2 Division of structures based on their proton loading ability

The observed proton pumping behavior shows if the energy difference between the loaded and unloaded protonation state, $\Delta\Delta G_{UL-L}^{Ee}$, is larger than ≈8.4 kcal/mol, the protein is fixed in the loaded state with 2 protons bound (Figure 3, Table 2-catigory 1). All of the 12 crystal structures are in this locked, loaded state. There are 6 structures that have a small fraction of microstates with a third proton. They have an average ionization state -2.99 and 0.18 integrated proton uptake (category 1'). Thus, although these structures have a high proton affinity, microstates with three bound protons and -2 charge are never really energetically accessible.

The largest group of structural clusters have $\Delta\Delta G_{UL-L}^{Ee}$ between ≈-2.9 to ≈8.4 kcal/mol (Category 2, 84 structures). Here the loaded and unloaded states are close enough in energy so their probability is tuned by changes in heme *b* and the BNC. These have an average ionization state between -4 and -3 and the integrated binding from 0.07 to 3.51. When $\Delta\Delta G_{UL-L}^{Ee}$ is between ≈2 to ≈4 kcal/mol the system is poised to be most sensitive to the heme *b* and BNC states and there is maximal proton loading.

When $\Delta\Delta G_{UL-L}^{Ee}$ is more negative than ≈-2.9 kcal/mol the system is fixed in the unloaded state with net charge -4 and one bound proton (category 3) or switches to load and unload protons between the -5 and -4 ionization states (category 4). Sixteen structures have a very low proton affinity where a fully deprotonated microstate with a charge of -5. The minimum, average charge for this group is -4.40 indicating that the -4 and -5 states are close in free energy, although there is not much proton loading. These structures are all from the trajectories that imposed a -5 PLS charge, which may over stabilize this fully deprotonated state (F_eMS0 or F_{+e}MS0 (Table 1)).

3.2.3 Change in $\Delta\Delta G_{UL\text{-}L}$ in the different reaction states

The pumping behavior depends on both the absolute value of $\Delta\Delta G_{UL\text{-}L}$ and as well as how this value changes between states that are expected to bind and release protons. Table 3 compares the range of $\Delta\Delta G^{Ee}_{UL-L}$, which has the most protons loaded and $\Delta\Delta G^{Fp}_{UL-L}$ which has the least for each proton loading category. $\Delta\Delta G_{UL\text{-}L}$ is always less positive in the F_P state, favoring proton release. However, in the locked loaded states (category 1) $\Delta\Delta G_{UL\text{-}L}$ remains positive, while in the locked unloaded states (category 3) it remains negative. As shown schematically at the top of Figure 3, only in the active, category 2, state does $\Delta\Delta G_{UL\text{-}L}$ move from being positive in the E_e state favoring loading and negative in the F_P state. Category 5 moves between the -5 and -4 states so $\Delta\Delta G_{UL\text{-}L}$, which explores the difference between the -4 and -3 states, strongly, favors the -4 state and so is negative.

Figure SI.5 compares $\Delta\Delta G^{Fp}_{UL-L}$ with $\Delta\Delta G^{Ee}_{UL-L}$, $\Delta\Delta G^{Fe}_{UL-L}$, and $\Delta\Delta G^{F+e}_{UL-L}$. The proton affinity in the E_e state is on average 5.4 kcal/mol more than in the F_p state for all structures, showing how the PLS proton binding energy that is decreased when a chemical proton is added to the BNC. With this shift in $\Delta\Delta G_{UL\text{-}L}$ it is clear why $\Delta\Delta G^{Ee}_{UL-L}$ between 2 and 4 kcal/mol is optimal for proton loading as with a shift of 5-6 kcal/mol the $\Delta\Delta G_{UL\text{-}L}$ will be above zero to bind a proton in the E_e state and then be lowered to below zero in the F_P state where the proton is released. The $\Delta\Delta G^{Fe}_{UL-L}$ favors proton binding by only ≈3.3 kcal/mol more than $\Delta\Delta G^{Fp}_{UL-L}$. Thus, while the F_e substate is the peak of proton binding amongst the three F redox substates, this is likely to have fewer protons loaded than in the E_e substate (Figure 2D). The F_{+e} state only favors proton binding by ≈1.3 kcal/mol more than $\Delta\Delta G^{Fp}_{UL-L}$, which represents the small increase in PLS proton affinity when an electron is transferred to heme *b*. This small change is why there is little proton uptake on this step (Figure 2).

Table 2. Summary of all investigated structures

Category	Behavior	# of clusters	# of frames	Average ionization	Integrated proton uptake	$\Delta\Delta G^{Ee}_{UL-L}$	$\Delta\Delta G^{Fp}_{UL-L}$
1	locked loaded (-3)	23	471	-3.00	0.00 ~ 0.05	8.51 ~ 13.76	3.59 ~ 8.23
1'	*loaded** (-3)	*6*	*187*	*-2.99 ~ -2.96*	*0.07 ~ 0.44*	*8.41 ~ 12.83*	*3.25 ~ 8.34*
2	Active (-4 to -3)	84	2257	-3.99 ~ -3.02	0.07 ~ 3.51	-3.38 ~ 8.38	-9.51 ~ 1.82
3	locked unloaded (-4)	7	153	-4.00	0.00 ~ 0.41	-4.41 ~ -2.06	-9.80 ~ -4.80
4	Active-empty (-5 to -4)	16	442	-4.40 ~ -4.00	0.15 ~ 1.98	-4.72~ -2.31	- 9.45 ~ -3.81

The MD Analysis program [62–64] analyzed the positions of the six PLS residues in the seven MD trajectories finding 124 clusters. MCCE simulations are carried out on one snapshot from each cluster and on 12 crystal structures. The 136 analyzed structures are divided into 5 categories with different proton uptake behavior as defined by the average ionization state of the PLS and the proton uptake behavior shown in Figure 3. # of clusters: How many of the 136 clusters are in this category. This is the number of structures analyzed in each category; # of frames: The total number of frames in this category in all of seven MD trajectories and crystal structures; The last four columns give the range of values for all structures. Average ionization: average net charge of the six-residues of the PLS along the 12-substate reaction cycle (dot Figure 3A); Integrated proton uptake is calculated by Equation 2 (Figure

3B); $\Delta\Delta G_{UL-L}^{Ee}$ and $\Delta\Delta G_{UL-L}^{Fp}$ are the energy difference between the lowest energy unloaded (charge -4) and loaded (-3) microstates, calculated with Equation 5, in the E_e or F_p substates respectively. The imposed E_e state generally leads to maximal proton loading while the F_p substate has the fewest protons loaded.

3.2.4 The intra PLS interactions lead to the changes in PLS proton affinity

The free energy difference between the lowest energy microstate with a charge of -4 and -3 ($\Delta\Delta G_{UL\text{-}L}$) in any particular imposed heme *b*/BNC state varies by close to 20 kcal/mol in the different structures (Figure SI.3). MCCE calculates this energy difference as the sum of the loss of solvation energy as a charge is moved from water into the protein interior, the interactions with the backbone amide dipoles and the interaction with other residues in the protein. What is most notable is that the interaction with the residues in the protein that do not include the PLS favor the -4 charge state relative to the state with a -3 charge by -19 ± 3 kcal/mol. The average charge of the protein (not including the PLS or the cofactors) is 10.8±0.85 in the MCCE calculations. Thus, the surrounding protein is not the main source of the variability of proton affinity of the various structure, but provides a background positive potential that stabilizes the negative PLS states found here. The sum of the energy from residues desolvation and their interaction with the backbone dipoles is near zero. Thus, it is the electrostatic interactions amongst the six PLS residues that varies by ≈30 kcal/mol between the structures that stabilize the -5 charge state to those locked in the loaded state.

3.3 Where are the protons in the PLS?

The PLS consists of six residues so can generate 64 (2^6) pronation microstates. The MCCE analysis shows the net charge of the PLS cluster switches from one proton loaded to two protons loaded, changing charge from a charge of -4 to -3. There are six possible protonation microstates with one proton and 15 microstates with two protons. There is only one state with no protons and a charge of -5. It should be noted that MCCE generates ≈1 million microstates which can have different polar proton positions and residue tautomers for the PLS residues in addition to the different protonation states. The conformational microstate with the lowest energy defines the energy for each protonation state.

Figure 4 and Table SI.2 summarize the distribution of the favored, distinct protonation microstates for the structures grouped into 5 categories by their proton loading behaviors (Table 2). Of the 64 possible protonation states, only those with a charge state of -5 to -2 have low enough energy (within 10 kcal/mol of the minimum) to have any microstate occupancy, so those with a net charge of -1, 0 and 1 are not shown. The vertical ordering of microstates within a given charge is random. The structures within a category are ordered along the x axis by the average PLS charge through the 12 imposed states for heme *b* and the BNC (dot in Figure 3A). The small black boxes show the highest probability, lowest energy microstate for that structure. Microstates with energies ≈3kcal/mol above the lowest energy state are orange, with yellow and white regions at still higher energy.

Five microstates that are found to be at lowest energy are schematically shown in Figure 5. The lowest energy microstates are almost always MS 14 and 16 with a charge of -3; and MS 2 and 5 with a charge of -4 and MS 0 with a charge of -5. Each of these have PRD_{a3}, Asp287 and Glu126B ionized, while the protonation of PRA_{a3}, Asp372 and His376 vary. Table SI.2 gives information about minority protonation distributions. These include microstates with -3 or -4 charge where PRD_{a3} or Glu126B are protonated, which are at higher but still accessible energy.

Figure 4A shows the E_e, maximally loaded and 4B the F_p, maximally unloaded states. When the lowest energy states are the same in the two figures, the system is locked. Almost all structures in the locked, loaded state (category 1) are in MS 16 with His376 and Asp372 protonated (Figure 5). Category 2, which can load and unload protons through the reaction cycle, shifts the position of the low energy states with MS 14 (Asp372 and PRA_{a3} protonated) as the predominate loaded species and MS 5 (Asp372 protonated) as the unloaded microstate. Category 3 which is fixed in the unloaded state or category 4, which can access the -5 net charge state tend to have PRA_{a3} protonated (MS 2). Thus, we see that structures that have locked protonation states (category 1 and 3) choose different proton positions than the active structures (category 2). So there is an active (MS 5) and a locked (MS 2) unloaded state as well as an active (MS 14) and locked (MS 16) loaded state.

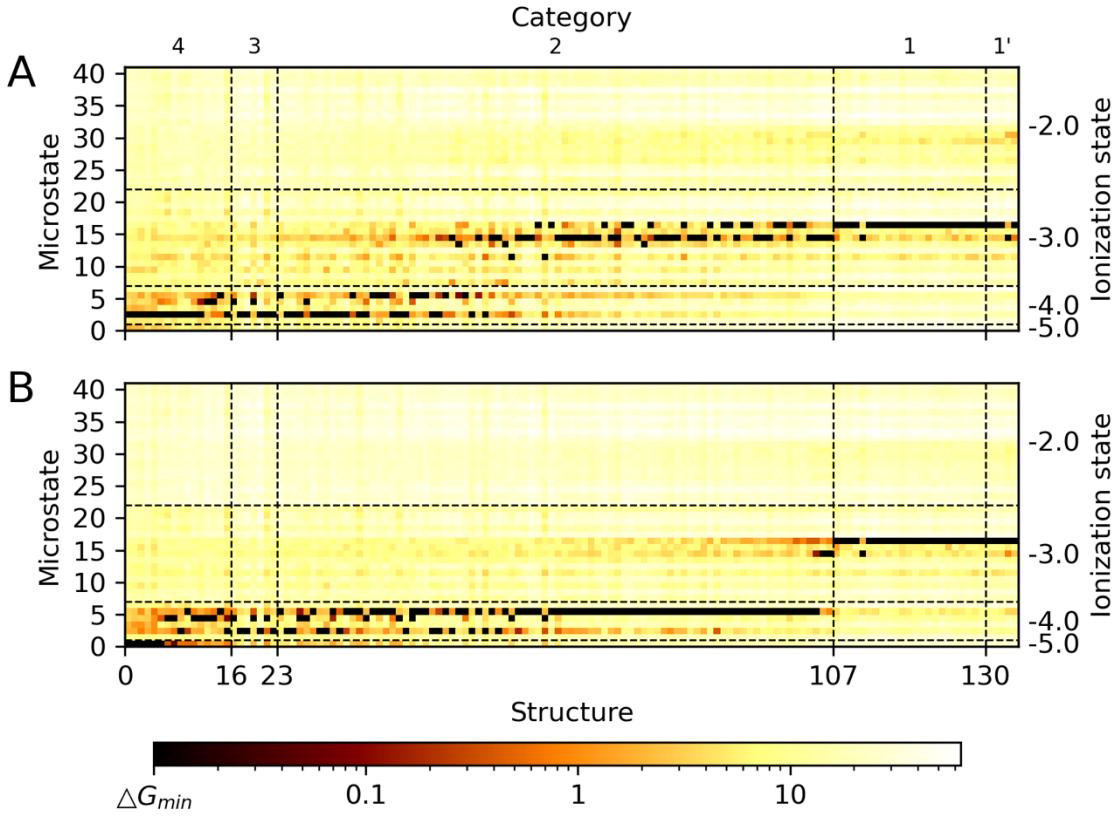

Figure 4 Energy map for (A) ΔG_{MS}^{Ee} and (B) ΔG_{MS}^{Fp}. The protonation microstates are numbered from 0 to 63 while the structures from 0 to 135. The microstate numbering is sorted by ionization state of the PLS ascending, from -5.0 (zero protons bound) to -2.0 (3 PLS protons). More positive PLS states are all at unattainable energy and are not shown. The structures are sorted along the x axis by the proton pumping categories (Table 2). Black boxes show microstates with lowest energy: MS 0, no proton on PLS; MS 2, a proton on PRA_{a3}; MS 5, a proton on Asp372; MS 14, a proton on Asp372 and PRA_{a3}; MS 16, a proton on Asp372 and His376 (Figure 5). Orange boxes: Less probable, but accessible microstates, which are listed in Table SI.2. The energy bar shows energy relative to the minimum energy for that structure in kcal/mol. Energies are shown on a log scale so that a wider range of energies can be visualized.

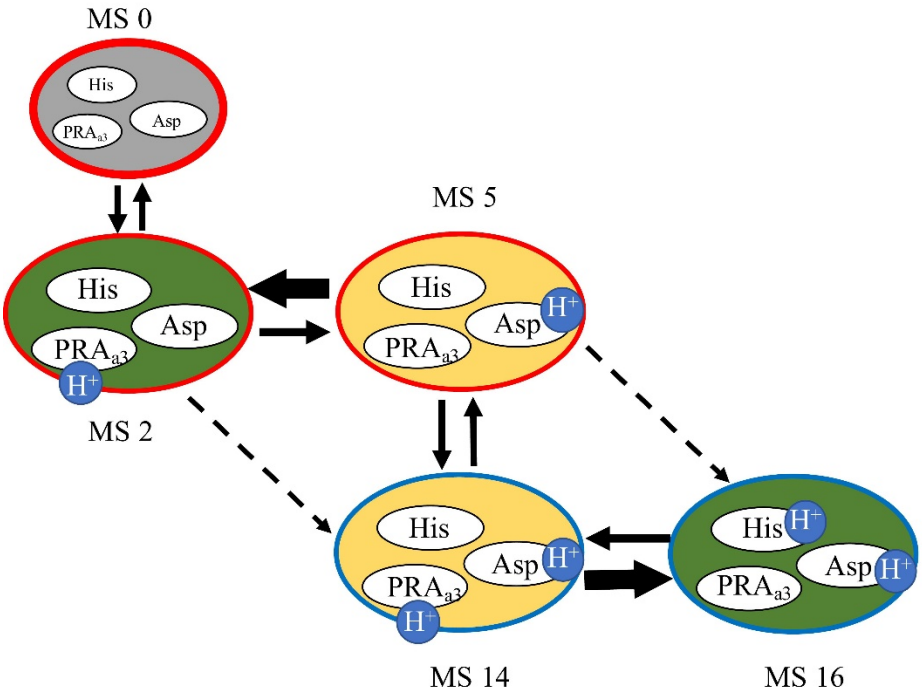

Figure 5 Lowest energy protonation microstates (Figure 4, Table SI.2). Protonation states of His376, Asp372 and the PRA_{a3} are shown. In these high probability states Asp 287, Glu126B and the PRD_{a3} are always ionized. MS numbering is the same as Figure 4. MS 14 and MS 16 have a PLS charge of -3; MS 2 and MS 5 have a charge of -4; and MS 0 a charge of -5. MS 16 is locked loaded; MS 14 and MS 5 are close in energy, so their population is modulated by the BNC and heme b charge in the reaction cycle. MS 2 can be locked unloaded or moved to MS 0 through the imposed reaction cycle. The orange filled ellipsoids represent active microstates, while green filled ellipsoids the locked states. The red outline of ellipsoid indicates the PLS is in an unloaded state while the blue outline indicates a loaded state. The MS 0 is shown in grey filled ellipsoid and a thicker red outline indicating a more negative state. The dashed arrows from MS 2 to MS 14 and MS 5 to MS 16 show highly unfavorable moves from a locked to an active microstate with a different number of protons.

3.4 Structure changes that lead to different proton pumping behavior

3.4.1 The protonation states assigned in the MD trajectory influence the proton loading behavior

The different input trajectories yield structures that favor different proton distributions leading to different proton uptake behavior (Table 3, the dot color in Figure 3 and Figure SI.6). Trajectories in the F_{+e} redox state (Table 1) were carried out in the five protonation microstates. The initial trajectories were run in the MS 0 state with no protons in the PLS and MS 2 and MS 5 with a proton on PRA_{a3} or Asp372. MCCE analysis, showed the importance of states with 2 protons. MS 14 was identified as an active loaded protonation states and while all crystal structures were found in the locked and loaded protonation state MS 16. Trajectories were also prepared in these states.

Figure SI.6 replots the energy map (Figure 4) sorted by the input trajectory rather than the PLS loading category. Some trajectories return only one protonation state, while others have a larger diversity of accessible protonation states in the Boltzmann distribution. The proton affinity (characterized by $\Delta\Delta G^{Ee}_{UL-L}$), is smallest in snapshots from trajectories with fewer protons so increases from F_{+e}MS0, F_{+e}MS2, F_{+e}MS5, F_{+e}MS14, to F_{+e}MS16. Connecting the trajectory protonation states with the resultant MCCE protonation states the proton affinity tends to be

lower in snapshots from (F_{+e}MS2) as MCCE finds MS 2 is the preferred protonation states in the locked unloaded trajectory. The F_{+e}MS5 trajectory, with the active, unloaded protonation state MS 5, leads to more active structures. Likewise, the proton affinity is lower in the trajectory (F_{+e}MS14) using the active loaded protonation state (MS 14) than the trajectory (F_{+e}MS16) using the locked loaded protonation state (MS 16). A key finding is that trajectories run with protonation states MCCE considers to be active (MS 5 and MS 14) return structures that can toggle between loaded and unloaded states in MCCE rather than being trapped in the protonation state of the trajectory. All crystal structures have a very high PLS proton affinity, and never release protons being locked, loaded state (MS 16) independent of the imposed heme *b* and BNC charges.

Trajectories were also run in different redox states. In the MS 2 protonation state, snapshots initiated from the F_e trajectory have a broader proton binding energy distribution than the F_{+e} trajectories. Structures from F_eMS0 have what we consider possibly artifactually low proton affinity so that microstates without protons are accessible.

Table 3. Dependence of the proton loading behavior on the trajectory redox and protonation state.

Trajectory	Trajectory Protonation	# of clusters	Integrated proton uptake	Average ionization	$\Delta\Delta G_{UL-L}^{Ee}$	$\Delta\Delta G_{UL-L}^{Fp}$	$\Delta\Delta\Delta G_{UL-L}^{Ee-Fp}$
F_eMS0	PRA_{a3}^-	17	0.65±0.66	-4.08±0.13	-2.78±1.08	-6.70±1.61	3.92
F_eMS2	$PRA_{a3}H$	24	0.96±1.01	-3.89±0.14	-0.57±2.67	-6.12±1.41	5.55
F_{+e}MS0	PRA_{a3}^-	21	0.60±0.52	-3.97±0.09	-1.91±1.41	-6.28±1.57	4.37
F_{+e}MS2	$PRA_{a3}H$	19	2.46±0.80	-3.63±0.19	2.38±1.70	-4.47±1.44	6.85
F_{+e}MS5	D372H	14	1.88±0.63	-3.39±0.24	3.43±1.25	-2.84±1.65	6.27
F_{+e}MS14	$PRA_{a3}H$, D372H	16	1.17±1.00	-3.15±0.16	6.42±2.43	-0.24±2.94	6.66
F_{+e}MS16	H376H, D372H	13	0.03±0.04	-3.00±0.00	11.03±1.13	6.08±1.22	4.95
crystal	Crystal	12	0.01±0.02	-3.00±0.00	11.88±1.66	6.84±1.57	5.04

The assigned redox and protonation states for the MD trajectory are given in Table 1. # of clusters: Meta-analysis divides the trajectory into clusters based on the positions of the six PLS residues. One structure from each cluster is analyzed. Integrated proton uptake is calculated with Equation 2, showing the proton uptake along the reaction cycle. The average ionization is the average net charge of the six-residue PLS in the 12-substate reaction cycle. $\Delta\Delta G_{UL-L}^{Ee}$ and $\Delta\Delta G_{UL-L}^{Fp}$ are the energy difference between the lowest energy state with a PLS charge of -4 and the one with a charge of -3 in the E_e and F_p substates (Equation 5). $\Delta\Delta\Delta G_{UL-L}^{Ee-Fp}$ is the difference of this proton affinity in the E_e and F_p substates.

3.4.2 Structural change in different proton pumping behaviors

Intra-cluster structural variations move the free energy difference between loaded and unloaded states of the PLS into an active region where $\Delta\Delta G_{UL-L}$ can cross from being below to above zero as the protein cycles between imposed loading or unloading redox/protonation substates of heme b and the BNC. In locked structures $\Delta\Delta G_{UL-L}$ is either too positive or too negative in all redox substates for the PLS to change protonation states. The distances between

key PLS residues are shown as a function of the proton loading category (Figure 6 left) or sorted by the input trajectory (Figure 6 right).

D287 and H376 sit near the top of the PLS cluster (Figure 1B). They never come close enough to form a hydrogen bond, but the distance between them becomes shorter, showing a more compact PLS, as the proton loaded states are favored (moving to the right in each panel in Figure 6). In MS 16, which is the locked and loaded state (category 1), H376 and D372 are protonated and both of these residues make hydrogen bonds with the ionized PRA_{a3}. The protonated H376 also makes a hydrogen bond with the deprotonated E126B. The presence of these 3 hydrogen bonds stabilizes the state with 2 protons into locked, loaded behavior. In the rare structures where the -2 states, with 3 protons, become possible (category 1') these hydrogen bonds become less stable.

In the active states (category 2) D372 retains its proton while H376 is deprotonated. The loaded MS 14 (PRA_{a3} protonated) or unloaded MS 5 are most likely to be populated. Only the hydrogen bond between the propionic acid and H376 is consistently made. The locked, unloaded structures (category 3, MS 2, PRA_{a3} protonated) often retain this one hydrogen bond and begin to develop a hydrogen bond between Asp372 and His376.

Snapshots that come from the MS 0 trajectories are prepared with no protons and a maximal charge of -5, which appears to introduce intra-PLS repulsion. The distance between D287 and H376 at the top of the PLS is approximately 2 Å further apart than in the structures that favor a -3 charge. PRA_{a3} is no longer hydrogen bonded to any of the other PLS residues. The neutral H376 and ionized D372 now often make a hydrogen bond.

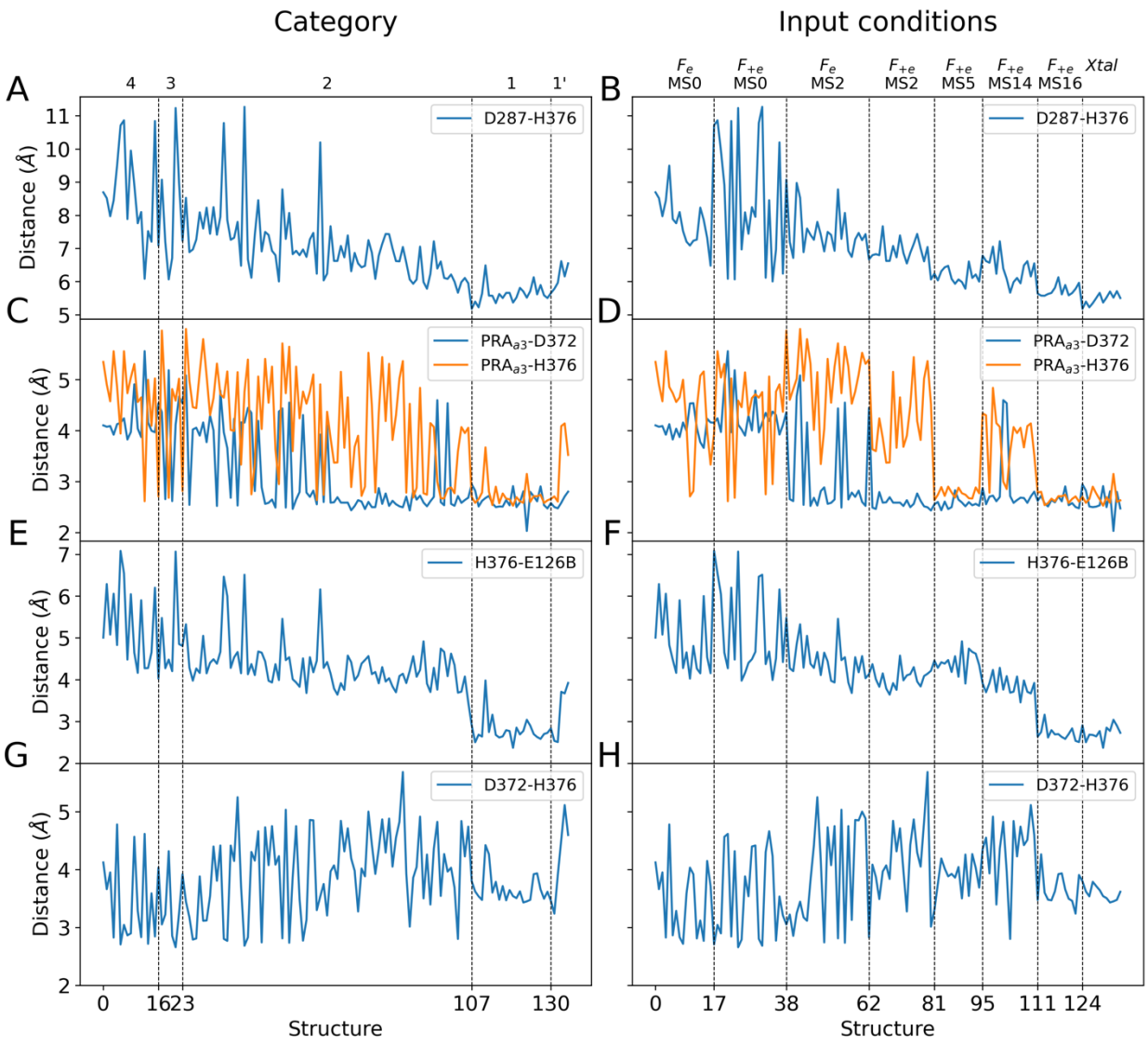

Figure 6 PLS structure changes. The x-axis is ordered by the average PLS charge (red dot in Figure 3A) in the left column. In the right column the structures are first ordered by the MD trajectory that is the source of the MCCE input file and then by the average PLS charge change within that catigory. Within each panel, structures locked in the unloaded state are on the left; the structures that can participate in active loading and unloading are in the middle and those locked in the loaded state are on the right. The distances changes between Asp287 and His376, PRA$_{a3}$ and Asp372, PRA$_{a3}$ and His376, His376 and Glu126B, Asp372 and His376 are shown individually. The distances plotted are the minimum distance between either side chain oxygen or side chain nitrogen. Fig 2A-C show examples of structures from category 4 (orange), 2 (purple) and 1 (red).

4. Discussion

4.1 proton loading/unloading mechanism of PLS

The pattern of proton loading to the PLS through the reaction cycle. A Proton Loading Site, PLS, is found within a hydrogen bonded cluster (Table SI.4), on the P-side of heme b and the BNC consisting of six groups (Figure 2A-2C). The MCCE calculations follow the protonation states of the PLS residues as a function of imposed changes to the redox and protonation states of heme *b* and the BNC (Figure 2D). The calculations propose the substates that lead to proton loading and unloading [9,72]. The PLS always has the most proton loaded when the electron is transferred from heme *b* to the BNC (X_e substates where X can be R, F, O, E (Table SI.1)). This step has been previously suggested to trigger proton loading [16]. After a proton has been added to the BNC, to form the X_p substates, the PLS proton affinity decreases, releasing the loaded proton. The net charge imposed on CcO in X_{+e} and X_e are the same (-1) but moving the electron from heme *b* to the BNC does not increase the proton affinity significantly. Addition an electron to the protein by reducing heme *b* (X_{+e}) is associated with modest, variable proton uptake (Figure 2D).

The number of protons in the loaded and unloaded PLS. With six residues there are 6 possible charge states, from -5, with no protons) to a state where all groups were protonated, which would have a charge of +5. The MCCE calculations on most structures support a model where an unloaded structure has one proton bound and the loaded structure has two. The large negative charge is supported by a very positive electrostatic potential from the rest of the protein, which is similar for all structures (Figure SI.3). States with three protons are significantly higher in energy and are rarely populated even when the cluster proton affinity is very high (very positive $\Delta\Delta G_{UL\text{-}L}$).

The "default" protonation gives residues the protonation state they would have in solution. This has a PLS with neutral His and ionized Asp, Glu and Propionic acids, thus no bound protons and so a charge of -5. This state had been suggested as the unloaded state [38–42,73]. A subset of the snapshots derived from trajectories carried out in this protonation state (MS 0) do support the unloaded state with a charge of -5 (Figure SI.6). The PLS in these structures tend to break key intra-PLS hydrogen bonds and the distance between D287 and H376 near the top of the PLS show an expansion of the cluster (Figure 6B). We suggest, but cannot prove, that the PLS with no protons is an artifact, with the MD simulation returning structures that stabilize the too negative input charge.

The $\Delta\Delta G_{UL\text{-}L}$, which is the difference between the lowest energy unloaded (charge -4) and loaded (charge -3) microstate in a given imposed substate of heme *b* and the BNC, explains much of the PLS proton loading behavior. Redox substates with higher proton affinity have more positive $\Delta\Delta G_{UL\text{-}L}$. $\Delta\Delta G_{UL\text{-}L}$ must move from being positive in imposed substates that load protons (e.g. X_e), to negative in substates that release protons (X_p) (schematic top Figure 2). The proton loading behavior is thus found to be controlled by the absolute value of the proton affinity, which depends on the input structure (Figure 2D, Table SI. 1, Figure SI. 1). It should be noted that $\Delta\Delta G_{UL\text{-}L}$ in one state provides a qualitative not quantitate picture of the loading behavior. When this value is close to zero then the cluster is near its effective pK_a and the changes in potential generated by heme b/BNC charges will change the protons bound in each structure. However, different structures feel the imposed charges differently (Fig SI.4). The take home message is that if the background proton affinity, estimated by $\Delta\Delta G_{UL\text{-}L}$, is too large or small the PLS is locked. When the energy difference between the loaded and unloaded states is

small the PLS will be active, but the stoichiometry of proton loading will then be sensitive to the details of each PLS structure.

As the largest difference between the proton affinity to the change in imposed charges at heme *b* and the BNC is ≈5 kcal/mol, between the E_e and F_p substates (Figure SI.5). When the positions of the PLS residues make the proton affinity in the E_e state higher than ≈5.5 kcal/mol the PLS is locked in the loaded state (Figure 3). The proton affinity does change through the reaction cycle; but not by enough to have a significant promotion of the PLS clusters lose a proton. Likewise, if the proton affinity is less than zero in the maximally loaded E_e state the PLS will never bind a proton. Trajectories run with no PLS protons can have very low proton affinity and so cycle between the -5 and -4 charge states (Figure 2D).

The PLS has a range of different proton loading behaviors. A real, dynamic protein will change its structure to accommodate changes in the PLS protonation state. This relaxation can help the PLS maintain the loaded or unloaded protonation state for sufficient time to remain in sync with the opening and closing of the proton transfer channels so that protons are bound from the N-side and released to the P-side. It is not (yet) possible to keep the conformational and protonation states fully at equilibrium. MCCE brings the protonation states to equilibrium with the imposed redox/BNC protonation states, but with a rigid backbone, while MD trajectories will change conformation to stabilize the input protonation state. By using meta-analysis to cluster the PLS conformation in the input trajectories 124 clusters were identified and a single structure from each was used as an input to MCCE (Table 3). These generated 5 categories of PLS behavior (Table 2). Thus, the PLS can be locked in the loaded state. However, 68% of the clusters derived from the trajectory are active so protons are loaded and unloaded through the reaction cycle.

It is inevitable that MD trajectories will change conformation to stabilize the input protonation state. MD trajectory run with those sites protonated (F_{+e}MS16) return structures that are locked in this loaded conformation. While structures from MD trajectory with protonated PRA_{a3} and Asp372 (MS 14) is in loaded state, retain the activity to lose a proton from PRA_{a3} and become MS 5, which is consistent with the proposed mechanism here. The observation that MS 14 and MS 16 were identified in trajectories with different input protonation states lends additional support to their viability in the reaction mechanism.

The protonation states are different in the locked and active states with the same charge. Thus, there are two forms for both loaded (PLS charge -3) and unloaded (-4) structures (Figure 5). The majority of the loaded state have a proton on Asp372. However, when the second proton on His376 the -4 charge state is inaccessible as $\Delta\Delta G_{UL\text{-}L}$ is too positive (MS 16); while when the proton is localized on the PRA_{a3} (MS 14) the state where this proton is lost is close in energy. Likewise, when the deprotonated state keeps the single proton on the propionic acid (MS 2), the structure is locked in the unloaded form, while when the lone proton is on the Asp372 (MS 5) a second proton can be bound. The crystal structures are always in the locked, loaded state so we assume they stabilize the proton on His376 and Asp372. Having a more compact PLS, with hydrogen bonds between PRA_{a3} and Asp372, PRA_{a3} and H376 as well as between His376 and Glu126B, favors the locked, loaded state. The loss of these hydrogen bonds and expansion of the PLS favors unloading.

Thus, the calculations suggest that CcO is a pump where the location of the proton in the PLS differentiates between states that are longer lived (locked) or more dynamic (active). This mechanism may be similar to activation and inactivation process in voltage gated ion channels,

which have conformational change when channels are active open to block the channel and transit to inactive state [74].

If there are transitions between locked and active protonation states for both loaded and unloaded structures (Figure 5), the protein must retain the ability to return the proton to the position where the PLS can change its net charge. The transition from the locked, loaded state (MS 16) to any unloaded states (MS 2 or 5) is uphill by 6 to 48 kcal/mol. However, in the structures that favor the locked MS 16 the shift in a proton to form the active MS 14 requires only 3.4±2.2 kcal/mol (Figure SI.8). This suggests the trapped MS 16 can return to the active MS 14 and then lose a proton. The ΔΔG separating the locked unloaded MS 2 from the active unloaded MS 5 is 3.5±4.1 kcal/mol, so again a pathway of thermal activation to the active loading state followed by proton bindings seems feasible.

MD trajectories run with the protonation state that MCCE finds to be active and unloaded (MS 5) present a slightly different picture. Snapshots derived from this trajectory remain active but can populate MS 16 when the BNC/heme *b* charges are changed. This trajectory returns two hydrogen bonds, between Asp372 and PRA_{a3}, His376 and PRA_{a3}. While the MS 16 MD trajectory, which is locked in the MS 16 loaded protonation state, adds a third hydrogen bond between His376 and Glu126B. The structure with two hydrogen bonds may reflect the active state. Thus, the MS 5 state would bind a proton to PRA_{a3} (MS 14) and then transfer it to His376 (MS 16). In addition, conformational changes in the MD trajectories with MS 14 protonation state, His376 rotation followed by Glu126B reorientation occurs in nanoseconds, showing a path for proton release to the P-side.

Possible sequence of proton movement. These MCCE calculations focus on the low energy protonation state so do not directly consider the proton exit pathway. However, changes in the position of His376 and Glu126B are seen in the MD trajectory with protonated PRA_{a3} and Asp372 (F_{+e}MS14) that may facilitate proton motion towards the P-side (Figure SI.9). Earlier simulations examined a possible role for His376 protonation/rotation [46] and here His376 moves away from PRA_{a3}, breaking the hydrogen bond in ≈19ns in the Fe_{+e}MS14 (active loaded state) MD trajectory. Glu126B rotates the two oxygen positions after around 30ns (Figure SI.9). This provides a path to move the proton from the PRA_{a3} in the active, loaded MS 14 to His376 in locked, loaded MS 16. The rotation of the protonated His moves it closer to Glu126B, which can rotate and release the proton to P-side.

The model of loaded and unloaded states was obtained here by MC sampling with a classical force field of all protonation states focusing on identifying the lowest energy states in a range of structures. The key protonation states are quite similar to those found in an earlier analysis using MD coupled with DFT and PMF calculations which focused on a pathway for proton transfers on the P-side of the *ba3* CcO [46]. There it was suggested that the PLS cluster starts from a protonated Asp372, which is similar to MS 5 state here. Then PRA_{a3} receives a proton from Asp372 (MS 2), coupling with the His376 rotation. Then a proton is loaded to Asp372 (MS 14), PRA_{a3} passes a proton to His376 (MS 16), the trapped, loaded state here. Given the complexity of the PLS, with many coupled residues it is interesting that two very different computational approaches are converging to a similar picture of the PLS.

4.2 Comparison of the *ba3* and *aa3* CcO

Comparison of the PLS in A and B-type CcO. A similar MCCE based study to identify the PLS region in the *aa3* CcO showed proton binding and release was mainly carried out by the propionic acid groups of heme *a* and heme a_3 [25]. A somewhat larger group of six residues are

calculated here to play a role in the PLS of *ba3* CcO. The previous study of the *aa3* CcO also showed some structures, such as the 1M56 crystal structure, exhibited active proton pumping along the redox state change of BNC, while others such as 2GSM, were locked in unloaded state [25]. None of the limited group of structures was trapped in the fixed loaded state. At the time those calculations were carried out, MCCE did not have the capability to save and bin the protonation states of all of the microstates and determine their state energies so it is not possible to make a complete comparison with the earlier simulations.

Low efficiency of proton pumping in B-type CcO. A long-standing question is why the ba_3 CcO pumps fewer protons than the aa_3 analog [73,75]. The source of the lower efficiency is not known. It may arise from weaker coupling between the BNC and PLS so protons are not pumped on each of the four reduction steps or each step could have a lower yield of pumped protons. Alternatively, it may be that wrong-way proton loading from the P-side is blocked more effectively in A-type than in B-type CcO. It is suggested that it may because of the lack of D channel [22] and subunit III or because B-type CcO is more ancient in evolution compared to A-type CcO [75–77]. Here we find no active structure loads/unloads more than 3.5 protons in the full cycle as active structures. This results from the need to keep the loaded and unloaded states (MS 5 and MS 14) close in energy so their equilibrium population can be modulated by the changes occurring at the relatively distant redox sites. In addition, it is seen that some structures do not contribute at all to proton loading or unloading as they are energetically locked in a loaded or unloaded state. The diversity of the stoichiometry of proton loading found here may help explain the low efficiency of pumping in *ba3* CcO. However, the model with auxiliary locked states found here can enhance proton pumping if the system can move efficiently from an active state with small driving force for (un)loading into the neighboring locked state, which fully stabilizes the loaded or unloaded state. This trapping mechanism could lead to maximal, 4 protons, pumping only if it has the appropriate rate constants for moving between locked and active states to stay in synchrony with the electron and proton transfers at the BNC.

Conservation of PLS residues. The PLS in the *ba3* and *aa3* CcOs are both in a highly interconnected cluster in a similar location in the structure [8,25]. The propionic acids of heme *a* or *b* and heme a_3 are present in both the *Rb. sphaeroides aa3* CcO and the *Thermus thermophilus ba3* CcO evaluated here. Several PLS residues are conserved including (*Rb. sphaeroides* numbering in parenthesis) Asp372 (Asp407), His376 (Mg ligand His411), but Glu126B, and Asp287 are not. Two nearby Arg, Arg449 (Arg481) and Arg450 (Arg482) help produce the positive electrostatic potential that supports the negative net charge of the PLS. A nearby Trp229 in *ba3* CcO corresponds to Trp280 in *aa3* CcO.

The importance of the residues proposed as the PLS can be evaluated by the degree of conservation and the sensitivity to mutation. The heme propionic acids (PRA_{a3} and PRD_{a3}) must always be present. His376 is 100% conserved, while Asp287 position is found in 64% of related sequence [78]. Asp372 is an Asp (31%) or Asn (67%). Glu126B is usually Met (89%) and rarely Glu (3%).

Several site-directed mutations of nearby residues, including Asp372, His376, Tyr133, Asn377, Glu126B, showed no significant loss on either catalytic or pumping activity [40]. However, the specific mutations Asp327Ile and His376Asn abolished or diminished proton pumping while retaining O_2 reduction, catalytic activity. Glu126B, in the PLS cluster was previously suggested to contribute proton pumping, but CcO activity is not sensitive to its mutation and it is poorly conserved [40]. The results of the mutations T312V/D372V and T312V/H376F suggested that T312 and H376 may be involved in the proton backflow system

where protons are bound from the P-side [40]. Asp287Asn shows both low respiratory and low proton pumping activities [40]. Thus, as has been seen previously the use of a cluster rather than a single residue may build in more flexibility to maintain function [8].

5. Conclusions

Proton pumps must choreograph changes in proton affinity and proton accessibility to the two sides of the membrane to be in synchrony with the electron transfer and chemical proton delivery that underlie the reaction cycle. Here we have investigated the proton transfer to the PLS in the B-type CcO. The PLS proton transfers occur far from the active site so the information from the BNC must be felt at long range. The reaction steps occur on the microseconds to millisecond time scale, so the change in proton affinity that load or unload the PLS must last this long. In the work reported here we show that changes in the reduction and protonation of the heme *b* and the BNC have a direct impact on the free energy difference between the protonated and unprotonated PLS of ≈5 kcal/mol. This is sufficient to lead to a maximum of 3.5 protons loading and unloading through the reaction cycle, but this is only seen for a subset of structures where the free energy difference between these states is in a narrow window. However, for a working CcO small changes in PLS proton affinity need to be protected from too easy proton loading/unloading. Thus, the range of integrated proton uptake of the proteins in actively loading structures is uncomfortably sensitive to the exact energy difference between the loaded and unloaded states (Figure 3).

The results show the protein may lock the loaded or unloaded states by shifting the proton. The active structures stabilize protons in the PLS: with a proton on Asp372 and PRA_{a3} in the loaded states and on the Asp372 in the unloaded state. Structures that have moved a proton in the loaded PLS proton from the PRA_{a3} to His376 have stabilized the loaded state; Likewise, a locked unloaded state can be generated by moving the single proton from the Asp372 to the PRA_{a3}. This is a plausible mechanism with a transition between active and locked states to allow the modest, long-range electrostatic changes from the BNC and heme *b* to lead to substantial and sufficiently stable loading and unloading.

The PLS structures that have a higher proton affinity are more compact. The distance between Asp287 and His376, near the P-side of the PLS, ranges from 5.5 Å to 11 Å. When the two residues are far apart, the PLS stays locked in the unloaded state. As they come closer together the PLS becomes active. When they are close, the PLS is fixed in the loaded state. Therefore, the Asp287-His376 pair may act as a switch in the P-side that opens when proton is loaded to the PLS and close when proton is locked at PLS. With a less compact PLS, the longer distance between His376 and PRA_{a3} or His376 and Glu126B enable PLS cluster to be active state to lose proton or gain proton. Once the His376 is hydrogen-bonded with both PRA_{a3} and Glu126B, the PLS cluster is completely locked at loaded state.

6. Acknowledgements

Research funded by grant number MCB-1519640 from the National Science Foundation. Q.C. acknowledges National Science Foundation for the grant number NSF-CHE-1829555.

7. References

[1] V.R.I. Kaila, M.I. Verkhovsky, M. Wikström, Proton-coupled electron transfer in cytochrome oxidase, Chem. Rev. 110 (2010) 7062–7081. https://doi.org/10.1021/cr1002003.

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

[3] J.K. Lanyi, Bacteriorhodopsin, Annu. Rev. Physiol. 66 (2004) 665–688. https://doi.org/10.1146/annurev.physiol.66.032102.150049.

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

[5] M. Wikström, K. Krab, V. Sharma, Oxygen Activation and Energy Conservation by Cytochrome c Oxidase, Chem. Rev. 118 (2018) 2469–2490. https://doi.org/10.1021/acs.chemrev.7b00664.

[6] M.R. Gunner, R. Koder, The design features cells use to build their transmembrane proton gradient, Phys. Biol. 14 (2017) 013001. https://doi.org/10.1088/1478-3975/14/1/013001.

[7] D. Kaur, X. Cai, U. Khaniya, Y. Zhang, J. Mao, M. Mandal, M.R. Gunner, Tracing the Pathways of Waters and Protons in Photosystem II and Cytochrome c Oxidase, Inorganics. 7 (2019) 14. https://doi.org/10.3390/inorganics7020014.

[8] X. Cai, K. Haider, J. Lu, S. Radic, C.Y. Son, Q. Cui, M.R. Gunner, Network analysis of a proposed exit pathway for protons to the P-side of cytochrome c oxidase, Biochim. Biophys. Acta. (2018). https://doi.org/10.1016/j.bbabio.2018.05.010.

[9] C. von Ballmoos, P. Adelroth, R.B. Gennis, P. Brzezinski, Proton transfer in ba_3 cytochrome c oxidase from *Thermus thermophilus*, Biochim. Biophys. Acta. 1817 (2012) 650–657. https://doi.org/10.1016/j.bbabio.2011.11.015.

[10] M. Wikström, Cytochrome c oxidase: 25 years of the elusive proton pump, Biochim. Biophys. Acta. 1655 (2004) 241–247. https://doi.org/10.1016/j.bbabio.2003.07.013.

[11] H.J. Lee, J. Reimann, Y. Huang, P. Adelroth, Functional proton transfer pathways in the heme-copper oxidase superfamily, Biochim. Biophys. Acta. 1817 (2012) 537–544. https://doi.org/10.1016/j.bbabio.2011.10.007.

[12] M.M. Pereira, M. Santana, M. Teixeira, A novel scenario for the evolution of haem-copper oxygen reductases, Biochim. Biophys. Acta. 1505 (2001) 185–208. https://doi.org/10.1016/s0005-2728(01)00169-4.

[13] J. Hemp, R.B. Gennis, Diversity of the heme-copper superfamily in archaea: insights from genomics and structural modeling, Results Probl Cell Differ. 45 (2008) 1–31. https://doi.org/10.1007/400_2007_046.

[14] Y.C. Kim, G. Hummer, Proton-pumping mechanism of cytochrome c oxidase: a kinetic master-equation approach, Biochim. Biophys. Acta. 1817 (2012) 526–536. https://doi.org/10.1016/j.bbabio.2011.09.004.

[15] P. Brzezinski, R.B. Gennis, Cytochrome c oxidase: exciting progress and remaining mysteries, J Bioenerg Biomembr. 40 (2008) 521–531. https://doi.org/10.1007/s10863-008-9181-7.

[16] M. Wikström, V. Sharma, Proton pumping by cytochrome c oxidase - A 40 year anniversary, Biochim. Biophys. Acta. 1859 (2018) 692–698. https://doi.org/10.1016/j.bbabio.2018.03.009.

[17] Y. Song, J. Mao, M.R. Gunner, Electrostatic environment of hemes in proteins: pK(a)s of hydroxyl ligands, Biochemistry. 45 (2006) 7949–7958. https://doi.org/10.1021/bi052182l.

[18] I. Belevich, M.I. Verkhovsky, Molecular mechanism of proton translocation by cytochrome c oxidase, Antioxid. Redox Signal. 10 (2008) 1–29. https://doi.org/10.1089/ars.2007.1705.

[19] D. Bloch, I. Belevich, A. Jasaitis, C. Ribacka, A. Puustinen, M.I. Verkhovsky, M. Wikström, The catalytic cycle of cytochrome c oxidase is not the sum of its two halves, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 529–533. https://doi.org/10.1073/pnas.0306036101.
[20] K. Faxén, G. Gilderson, P. Ädelroth, P. Brzezinski, A mechanistic principle for proton pumping by cytochrome c oxidase, Nature. 437 (2005) 286–289. https://doi.org/10.1038/nature03921.
[21] A. Kannt, T. Soulimane, G. Buse, A. Becker, E. Bamberg, H. Michel, Electrical current generation and proton pumping catalyzed by the ba3-type cytochrome c oxidase from Thermus thermophilus, FEBS Lett. 434 (1998) 17–22.
[22] H. Han, J. Hemp, L.A. Pace, H. Ouyang, K. Ganesan, J.H. Roh, F. Daldal, S.R. Blanke, R.B. Gennis, Adaptation of aerobic respiration to low O_2 environments, Proc. Natl. Acad. Sci. U.S.A. 108 (2011) 14109–14114. https://doi.org/10.1073/pnas.1018958108.
[23] E. Arslan, A. Kannt, L. Thöny-Meyer, H. Hennecke, The symbiotically essential cbb 3-type oxidase of Bradyrhizobium japonicum is a proton pump, FEBS Letters. 470 (2000) 7–10. https://doi.org/10.1016/S0014-5793(00)01277-1.
[24] G. Brändén, R.B. Gennis, P. Brzezinski, Transmembrane proton translocation by cytochrome c oxidase, Biochim. Biophys. Acta. 1757 (2006) 1052–1063. https://doi.org/10.1016/j.bbabio.2006.05.020.
[25] J. Lu, M.R. Gunner, Characterizing the proton loading site in cytochrome c oxidase, Proc. Natl. Acad. Sci. U.S.A. 111 (2014) 12414–12419. https://doi.org/10.1073/pnas.1407187111.
[26] M.R.A. Blomberg, The mechanism for oxygen reduction in the C family cbb3 cytochrome c oxidases - Implications for the proton pumping stoichiometry, J. Inorg. Biochem. 203 (2020) 110866. https://doi.org/10.1016/j.jinorgbio.2019.110866.
[27] P. Goyal, J. Lu, S. Yang, M.R. Gunner, Q. Cui, Changing hydration level in an internal cavity modulates the proton affinity of a key glutamate in cytochrome c oxidase, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 18886–18891. https://doi.org/10.1073/pnas.1313908110.
[28] I. Belevich, D.A. Bloch, N. Belevich, M. Wikström, M.I. Verkhovsky, Exploring the proton pump mechanism of cytochrome c oxidase in real time, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 2685–2690. https://doi.org/10.1073/pnas.0608794104.
[29] R. Sugitani, E.S. Medvedev, A.A. Stuchebrukhov, Theoretical and computational analysis of the membrane potential generated by cytochrome c oxidase upon single electron injection into the enzyme, Biochim. Biophys. Acta. 1777 (2008) 1129–1139. https://doi.org/10.1016/j.bbabio.2008.05.006.
[30] M. Wikström, M.I. Verkhovsky, Mechanism and energetics of proton translocation by the respiratory heme-copper oxidases, Biochim. Biophys. Acta. 1767 (2007) 1200–1214. https://doi.org/10.1016/j.bbabio.2007.06.008.
[31] S. Supekar, A.P. Gamiz-Hernandez, V.R.I. Kaila, A Protonated Water Cluster as a Transient Proton-Loading Site in Cytochrome c Oxidase, Angew. Chem. Int. Ed. Engl. 55 (2016) 11940–11944. https://doi.org/10.1002/anie.201603606.
[32] V.R.I. Kaila, V. Sharma, M. Wikström, The identity of the transient proton loading site of the proton-pumping mechanism of cytochrome c oxidase, Biochim. Biophys. Acta. 1807 (2011) 80–84. https://doi.org/10.1016/j.bbabio.2010.08.014.
[33] T. Yamashita, G.A. Voth, Insights into the mechanism of proton transport in cytochrome c oxidase, J. Am. Chem. Soc. 134 (2012) 1147–1152. https://doi.org/10.1021/ja209176e.
[34] A. Maréchal, B. Meunier, D. Lee, C. Orengo, P.R. Rich, Yeast cytochrome c oxidase: A model system to study mitochondrial forms of the haem–copper oxidase superfamily, Biochim Biophys Acta. 1817 (2012) 620–628. https://doi.org/10.1016/j.bbabio.2011.08.011.

[35] M.L. Björck, J. Vilhjálmsdóttir, A.M. Hartley, B. Meunier, L. Näsvik Öjemyr, A. Maréchal, P. Brzezinski, Proton-transfer pathways in the mitochondrial S. cerevisiae cytochrome c oxidase, Scientific Reports. 9 (2019) 1–8. https://doi.org/10.1038/s41598-019-56648-9.
[36] J. Quenneville, D.M. Popović, A.A. Stuchebrukhov, Combined DFT and electrostatics study of the proton pumping mechanism in cytochrome c oxidase, Biochim. Biophys. Acta. 1757 (2006) 1035–1046. https://doi.org/10.1016/j.bbabio.2005.12.003.
[37] J. Quenneville, D.M. Popović, A.A. Stuchebrukhov, Redox-Dependent pKa of CuB Histidine Ligand in Cytochrome c Oxidase, J. Phys. Chem. B. 108 (2004) 18383–18389. https://doi.org/10.1021/jp0467797.
[38] T. Soulimane, G. Buse, G.P. Bourenkov, H.D. Bartunik, R. Huber, M.E. Than, Structure and mechanism of the aberrant ba_3-cytochrome c oxidase from *Thermus thermophilus*, The EMBO Journal. 19 (2000) 1766–1776. https://doi.org/10.1093/emboj/19.8.1766.
[39] H.-Y. Chang, J. Hemp, Y. Chen, J.A. Fee, R.B. Gennis, The cytochrome ba_3 oxygen reductase from *Thermus thermophilus* uses a single input channel for proton delivery to the active site and for proton pumping, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 16169–16173. https://doi.org/10.1073/pnas.0905264106.
[40] H.-Y. Chang, S.K. Choi, A.S. Vakkasoglu, Y. Chen, J. Hemp, J.A. Fee, R.B. Gennis, Exploring the proton pump and exit pathway for pumped protons in cytochrome ba_3 from Thermus thermophilus, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 5259–5264. https://doi.org/10.1073/pnas.1107345109.
[41] J.A. Fee, D.A. Case, L. Noodleman, Toward a chemical mechanism of proton pumping by the B-type cytochrome c oxidases: application of density functional theory to cytochrome ba_3 of *Thermus thermophilus*, J. Am. Chem. Soc. 130 (2008) 15002–15021. https://doi.org/10.1021/ja803112w.
[42] C. Koutsoupakis, T. Soulimane, C. Varotsis, Probing the Q-proton pathway of ba3-cytochrome c oxidase by time-resolved Fourier transform infrared spectroscopy, Biophys. J. 86 (2004) 2438–2444. https://doi.org/10.1016/S0006-3495(04)74300-3.
[43] C. Koutsoupakis, T. Soulimane, C. Varotsis, Ligand binding in a docking site of cytochrome C oxidase: a time-resolved step-scan Fourier transform infrared study, J. Am. Chem. Soc. 125 (2003) 14728–14732. https://doi.org/10.1021/ja036107e.
[44] V. Daskalakis, S.C. Farantos, V. Guallar, C. Varotsis, Regulation of Electron and Proton Transfer by the Protein Matrix of Cytochrome c Oxidase, J. Phys. Chem. B. 115 (2011) 3648–3655. https://doi.org/10.1021/jp1115993.
[45] W.-G. Han Du, A.W. Götz, L. Noodleman, A Water Dimer Shift Activates a Proton Pumping Pathway in the PR → F Transition of ba3 Cytochrome c Oxidase, Inorg Chem. 57 (2018) 1048–1059. https://doi.org/10.1021/acs.inorgchem.7b02461.
[46] L. Yang, Å.A. Skjevik, W.-G. Han Du, L. Noodleman, R.C. Walker, A.W. Götz, Water exit pathways and proton pumping mechanism in B-type cytochrome c oxidase from molecular dynamics simulations, Biochim. Biophys. Acta. 1857 (2016) 1594–1606. https://doi.org/10.1016/j.bbabio.2016.06.005.
[47] I. Belevich, M.I. Verkhovsky, M. Wikström, Proton-coupled electron transfer drives the proton pump of cytochrome c oxidase, Nature. 440 (2006) 829–832. https://doi.org/10.1038/nature04619.
[48] M. Wikström, M.I. Verkhovsky, G. Hummer, Water-gated mechanism of proton translocation by cytochrome c oxidase, Biochim. Biophys. Acta. 1604 (2003) 61–65.
[49] J.E. Morgan, M.I. Verkhovsky, M. Wikström, The histidine cycle: a new model for proton translocation in the respiratory heme-copper oxidases, J. Bioenerg. Biomembr. 26 (1994) 599–608. https://doi.org/10.1007/bf00831534.
[50] P.R. Rich, Towards an Understanding of the Chemistry of Oxygen Reduction, Functional Plant Biol. 22 (1995) 479–486. https://doi.org/10.1071/pp9950479.

[51] D.M. Popović, A.A. Stuchebrukhov, Proton pumping mechanism and catalytic cycle of cytochrome c oxidase: Coulomb pump model with kinetic gating, FEBS Letters. 566 (2004) 126–130. https://doi.org/10.1016/j.febslet.2004.04.016.
[52] S.A. Siletsky, A.S. Pawate, K. Weiss, R.B. Gennis, A.A. Konstantinov, Transmembrane Charge Separation during the Ferryl-oxo → Oxidized Transition in a Nonpumping Mutant of Cytochrome c Oxidase, J. Biol. Chem. 279 (2004) 52558–52565. https://doi.org/10.1074/jbc.M407549200.
[53] M. Wikström, M.I. Verkhovsky, Towards the mechanism of proton pumping by the haem-copper oxidases, Biochim. Biophys. Acta. 1757 (2006) 1047–1051. https://doi.org/10.1016/j.bbabio.2006.01.010.
[54] H. Michel, Cytochrome c Oxidase: Catalytic Cycle and Mechanisms of Proton Pumping-A Discussion, Biochemistry. 38 (1999) 15129–15140. https://doi.org/10.1021/bi9910934.
[55] A.L. Woelke, A. Wagner, G. Galstyan, T. Meyer, E.-W. Knapp, Proton Transfer in the K-Channel Analog of B-Type Cytochrome c Oxidase from Thermus thermophilus, Biophys J. 107 (2014) 2177–2184. https://doi.org/10.1016/j.bpj.2014.09.010.
[56] T. Tiefenbrunn, W. Liu, Y. Chen, V. Katritch, C.D. Stout, J.A. Fee, V. Cherezov, High Resolution Structure of the ba3 Cytochrome c Oxidase from Thermus thermophilus in a Lipidic Environment, PLOS ONE. 6 (2011) e22348. https://doi.org/10.1371/journal.pone.0022348.
[57] C.Y. Son, A. Yethiraj, Q. Cui, Cavity hydration dynamics in cytochrome c oxidase and functional implications, Proc. Natl. Acad. Sci. U.S.A. 114 (2017) E8830–E8836. https://doi.org/10.1073/pnas.1707922114.
[58] L.M. Hunsicker-Wang, R.L. Pacoma, Y. Chen, J.A. Fee, C.D. Stout, A novel cryoprotection scheme for enhancing the diffraction of crystals of recombinant cytochrome ba3 oxidase from Thermus thermophilus, Acta Cryst D. 61 (2005) 340–343. https://doi.org/10.1107/S0907444904033906.
[59] B. Liu, Y. Chen, T. Doukov, S.M. Soltis, C.D. Stout, J.A. Fee, Combined Microspectrophotometric and Crystallographic Examination of Chemically Reduced and X-ray Radiation-Reduced Forms of Cytochrome ba3 Oxidase from Thermus thermophilus: Structure of the Reduced Form of the Enzyme, Biochemistry. 48 (2009) 820–826. https://doi.org/10.1021/bi801759a.
[60] V.M. Luna, J.A. Fee, A.A. Deniz, C.D. Stout, Mobility of Xe Atoms within the Oxygen Diffusion Channel of Cytochrome ba3 Oxidase, Biochemistry. 51 (2012) 4669–4676. https://doi.org/10.1021/bi3003988.
[61] W. McDonald, C. Funatogawa, Y. Li, I. Szundi, Y. Chen, J.A. Fee, C.D. Stout, Ó. Einarsdóttir, Ligand Access to the Active Site in *Thermus thermophilus* ba_3 and Bovine Heart aa_3 Cytochrome Oxidases, Biochemistry. 52 (2013) 640–652. https://doi.org/10.1021/bi301358a.
[62] R.J. Gowers, M. Linke, J. Barnoud, T.J.E. Reddy, M.N. Melo, S.L. Seyler, J. Domański, D.L. Dotson, S. Buchoux, I.M. Kenney, O. Beckstein, MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations, Proceedings of the 15th Python in Science Conference. (2016) 98–105. https://doi.org/10.25080/Majora-629e541a-00e.
[63] N. Michaud-Agrawal, E.J. Denning, T.B. Woolf, O. Beckstein, MDAnalysis: a toolkit for the analysis of molecular dynamics simulations, J Comput Chem. 32 (2011) 2319–2327. https://doi.org/10.1002/jcc.21787.
[64] M. Tiberti, E. Papaleo, T. Bengtsen, W. Boomsma, K. Lindorff-Larsen, ENCORE: Software for Quantitative Ensemble Comparison, PLOS Computational Biology. 11 (2015) e1004415. https://doi.org/10.1371/journal.pcbi.1004415.
[65] Y. Song, J. Mao, M.R. Gunner, MCCE2: improving protein pKa calculations with extensive side chain rotamer sampling, J Comput Chem. 30 (2009) 2231–2247. https://doi.org/10.1002/jcc.21222.
[66] M.R. Gunner, N.A. Baker, Continuum Electrostatics Approaches to Calculating pKas and Ems in Proteins, Meth. Enzymol. 578 (2016) 1–20. https://doi.org/10.1016/bs.mie.2016.05.052.

[67] J. Zhang, M.R. Gunner, Multiconformation continuum electrostatics analysis of the effects of a buried Asp introduced near heme a in *Rhodobacter sphaeroides* cytochrome c oxidase, Biochemistry. 49 (2010) 8043–8052. https://doi.org/10.1021/bi100663u.
[68] Y. Song, E. Michonova-Alexova, M.R. Gunner, Calculated proton uptake on anaerobic reduction of cytochrome C oxidase: is the reaction electroneutral?, Biochemistry. 45 (2006) 7959–7975. https://doi.org/10.1021/bi052183d.
[69] P.E.M. Siegbahn, M.R.A. Blomberg, Energy diagrams and mechanism for proton pumping in cytochrome c oxidase, Biochim. Biophys. Acta. 1767 (2007) 1143–1156. https://doi.org/10.1016/j.bbabio.2007.06.009.
[70] R. Liang, J.M.J. Swanson, Y. Peng, M. Wikström, G.A. Voth, Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome c oxidase, Proc. Natl. Acad. Sci. U.S.A. 113 (2016) 7420–7425. https://doi.org/10.1073/pnas.1601982113.
[71] H.J. Lee, L. Ojemyr, A. Vakkasoglu, P. Brzezinski, R.B. Gennis, Properties of Arg481 mutants of the aa3-type cytochrome c oxidase from Rhodobacter sphaeroides suggest that neither R481 nor the nearby D-propionate of heme a3 is likely to be the proton loading site of the proton pump, Biochemistry. 48 (2009) 7123–7131. https://doi.org/10.1021/bi901015d.
[72] C. von Ballmoos, P. Lachmann, R.B. Gennis, P. Ädelroth, P. Brzezinski, Timing of Electron and Proton Transfer in the ba_3 Cytochrome c Oxidase from *Thermus thermophilus*, Biochemistry. 51 (2012) 4507–4517. https://doi.org/10.1021/bi300132t.
[73] C. von Ballmoos, N. Gonska, P. Lachmann, R.B. Gennis, P. Ädelroth, P. Brzezinski, Mutation of a single residue in the ba_3 oxidase specifically impairs protonation of the pump site, Proc. Natl. Acad. Sci. U.S.A. 112 (2015) 3397–3402. https://doi.org/10.1073/pnas.1422434112.
[74] R.W. Aldrich, Fifty years of inactivation, Nature. 411 (2001) 643–644. https://doi.org/10.1038/35079705.
[75] V. Rauhamäki, M. Wikström, The causes of reduced proton-pumping efficiency in type B and C respiratory heme-copper oxidases, and in some mutated variants of type A, Biochim. Biophys. Acta. 1837 (2014) 999–1003. https://doi.org/10.1016/j.bbabio.2014.02.020.
[76] V. Sharma, M. Wikström, A structural and functional perspective on the evolution of the heme–copper oxidases, FEBS Letters. 588 (2014) 3787–3792. https://doi.org/10.1016/j.febslet.2014.09.020.
[77] C. Brochier-Armanet, E. Talla, S. Gribaldo, The Multiple Evolutionary Histories of Dioxygen Reductases: Implications for the Origin and Evolution of Aerobic Respiration, Mol Biol Evol. 26 (2009) 285–297. https://doi.org/10.1093/molbev/msn246.
[78] R. Schneider, A. de Daruvar, C. Sander, The HSSP database of protein structure-sequence alignments., Nucleic Acids Res. 25 (1997) 226–230.