STUDYING PROTON PUMPING MECHANISM OF BACTERIORHODOPSIN AND CYTOCHROME C OXIDASE WITH MULTI-CONFORMATION CONTINUUM ELECTROSTATICS

by

YIFAN SONG

A dissertation submitted to the Graduate Faculty in Physics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2006

UMI Number: 3232000

Copyright 2006 by
Song, Yifan

All rights reserved.

UMI Microform 3232000
Copyright 2006 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2006

YIFAN SONG

All Rights Reserved

This manuscript has been read and accepted for the
Graduate Faculty in Physics in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Marilyn R. Gunner

Date — Chair of Examining Committee

Sultan Catto

Date — Executive Officer

Timothy H. Boyer
David Frick
Thomas Haines
Ronald Koder
Themis Lazaridis
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

STUDYING PROTON PUMPING MECHANISM OF BACTERIORHODOPSIN AND CYTOCHROME C OXIDASE WITH MULTI-CONFORMATION CONTINUUM ELECTROSTATICS

by

YIFAN SONG

Advisor: Professor Marilyn R. Gunner

The proton gradient across the biological membrane is important for the biological systems. Bacteriorhodopsin and cytochrome c oxidase convert different energy sources into this gradient. The focus of this thesis is to understand the mechanism of these proteins using computational methods.

In bacteriorhodopsin, residue ionization states were calculated in 9 crystal structures trapped in bR, early M and late M states by Multi-Conformation Continuum Electrostatics (MCC). The three groups in the central cluster are ionized in bR structures while a proton has transferred from the SB^+ to Asp 85^- in the late M structures matching prior experimental results. The proton release cluster binds one proton in bR structure which is lost to water by pH 8 in late M. Modest changes in intra-protein interactions cause the charge shifts within the clusters. Motions of Arg 82 couple the proton shift in the central cluster to proton release. Changes in the total charge of the two clusters are coupled by direct long-range interactions.

Cytochrome *c* oxidase is a transmembrane proton pump that builds an electrochemical gradient using chemical energy from the reduction of O_2. Ionization states of all residues were calculated with MCCE in seven anaerobic oxidase redox states ranging from fully oxidized to fully reduced in *Rb. sphaeroides* cytochrome c oxidase. At pH 7, only a hydroxide coordinated to Cu_B shifts its pK_a from below 7 to above 7, and so picks up a proton when Heme a_3 and Cu_B are reduced. Glu I–286, Tyr I–288, His I–334 and a second hydroxide on Heme a_3 all have pK_as above 7. The propionic acids near the BNC are deprotonated with pK_as well below 7. This suggests electroneutrality in the BNC is not maintained during the anaerobic reduction. The electrochemical midpoint potential (E_m) of Heme a is calculated to shift down when the BNC is reduced, which agrees with prior experiments. If the BNC reduction is electroneutral, then the Heme a E_m is independent of the BNC redox state.

Acknowledgements

I would like to thank all of the people who have helped and supported me during my six years as a Ph.D. candidate. I would especially like to express my deep gratitude to:

Marilyn R. Gunner, my Ph.D. advisor, for her bringing me into the field of biophysics, for her encouragement and support, and for the advice she gave me, which has aided me not only in science, but also in how to become a better academic, and a better person.

Thomas Haines, for his guidance, support, and for our many helpful discussions.

Junjun Mao, for sharing your knowledge in computer programming and chemistry, and for the time we worked together programming MCCE2.

Gail Schneider, for carefully reading this dissertation and my other manuscripts.

Karin Hauser, Jinrang Kim, and Theodore W. Leslie, for all the time we spent together in the lab and elsewhere, for broadening my view in science, for their friendship through the years, and for exploring this wonderful city with me.

Anthony Schultz, for the wonderful time we spent together in class, and for being a great friend.

I am also especially grateful to:

My parents, for always supporting me.

New York, September 9th, 2006

Table of Contents

List of Tables

List of Illustrations

1. Introduction

1.1 Proton transfer in membrane proteins

Energy and matter are the two physical essences of life. Over billions of years, life has evolved to use different enzymes to extract energy from various sources. One basic mechanism uses a circuit formed of protons, electrons, and substrates through proteins in the cell or mitochondrial membrane (Fig 1.1). This circuit consists of two parts. In the first part, the energy from the environment is converted into electrochemical energy in the form of a proton gradient across the membrane (1). In the second part, the proton gradient is used to produce ATP (2,3), where energy is stored in the form of chemical bonds and can now be transported in the solution (4,5).

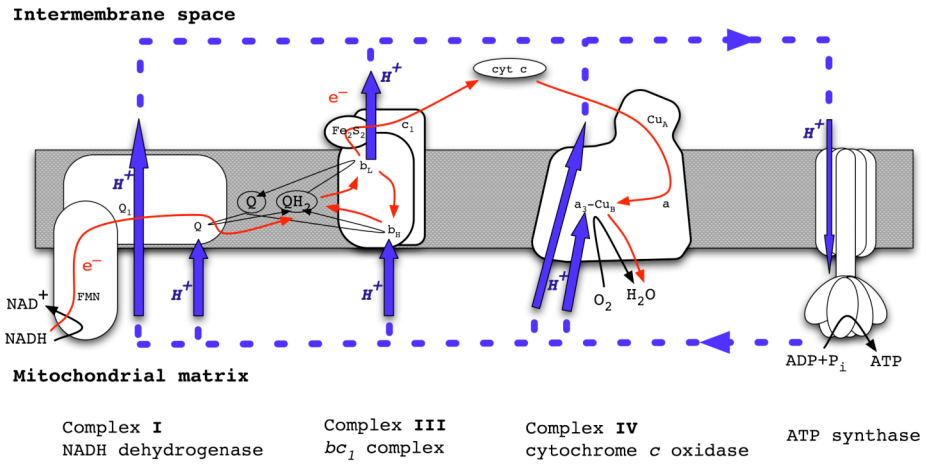

Figure 1.1 Schematic view of the respiratory chain in the inner mitochondrial membrane

Complex I, III and IV and ATP synthase from the respiratory chain are shown. Blue arrow: flux of protons; red arrow: flux of electrons

When life first arose on the earth, there was little dioxygen in the atmosphere (6,7). Photosynthetic reactions evolved, first in bateria (8), to use light, the most abundant energy source on the earth, to build a trans-membrane proton gradient. Various enzymes are involved in this type of reaction.

.

Some are very simple, such as bateriorhodopsin (9,10), which consists of only 248 amino acids and moves one proton across the membrane with each photon (11,12). Some have much more complex structures and functions. For example, photosystem II (PSII) consists of 25 subunits and several thousand amino acids (13). Not only is it involved in proton transfer across the membrane, it also uses water as a terminal electron donor and produces dioxygen as a byproduct (14,15).

As more and more dioxygen was released into the earth atmosphere by PSII, new types of reactions evolved to utilize this energy source. Dioxygen can be reduced back to H_2O. Protein evolved to convert the energy released from breaking the O-O bond into a transmembrane proton gradient. One class of proteins that carry out this reaction is the heme-copper oxidases (16,17).

Both bacteriorhodopsin and cytochrome c oxidase are proton pumps. They take protons from the low potential side of the membrane (low proton concentration, high pH), move them over 30Å across the membrane, and release them to the high potential side (high proton concentration, low pH). The ability of a proton pump to move protons in one direction efficiently depends on the concerted changes of charge state of a few key residues. Identifying these key groups and understanding the pumping mechanism is still a very challenging problem for biochemists and biophysicists. How do they trap the energy released from a chemical process within pico seconds so that it can be used to drive a reaction on the time scale of microseconds to seconds? How do these proteins change the proton affinity of each key group to bind and release a proton at different stages of the reaction? How do they control the proton transfer direction of each step so that the proton pump doesn't short-circuit? Most of the work presented here is devoted to understanding the proton pumping mechanism used by bacteriorhodopsin and cytochrome c oxidase.

.

1.2 Bacteriorhodopsin

Bacteriorhodopsin from the archaeon *Halobacterium salinarum* absorb photons with retinal chromophore (18-22). Light absorption leads to retinal isomerization, which initiates protein structural changes. These changes somehow control multiple-step internal proton transfers ensuring that proton uptake is from the high pH and proton release to the low pH side of the protein.

Bacteriorhodopsin is a helical bundle with seven trans-membrane helices (Fig 1.2). The chromophore, retinal, is covalently bound to Lys 216 at the center of the protein to form a Schiff base. Unlike ion and water channels, proton pumps don't have a pore through the protein. Instead, a proton transfer pathway is buried in the protein. Nine ionizable residues can be found in the trans-membrane region, the Schiff base, Asp 85, 96, 115 and 212, Glu 194 and 204, and Arg 82 and 134 (Fig 1.2). Asp 85 and 212 can both serve as a negative counterion of the Schiff base. These three residues form a cluster at the center of the protein (central cluster). Asp 96 sits on the cytoplasmic side, near the edge of the membrane. Glu 194 and 204 form another cluster (proton release cluster) near the extracellular edge of the membrane. Asp 115 is deeply buried in the trans-membrane region, on the cytoplasmic side of the Schiff base, but ten Ångstrom away from the proton transfer pathway. Arg 82 sits between the central cluster and the proton release cluster, and Arg 134 is near the proton release cluster.

.

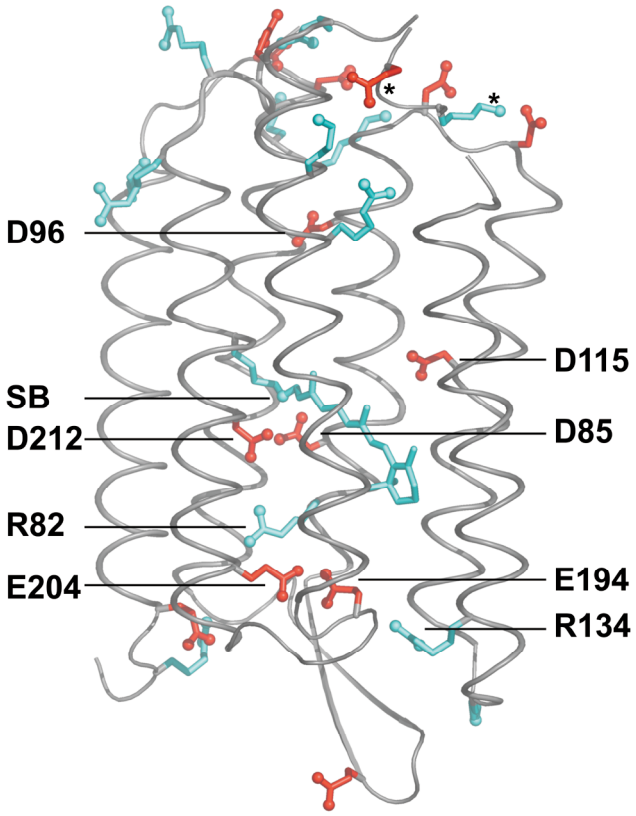

Figure 1.2 Ionizable residues in bacteriorhodopsin

Bacteriorhodopsin structure 1C8R. Acidic (red) and basic (blue) sidechains with terminal nitrogens or oxygens accented as spheres. All deeply buried residues are labeled. The retinal Schiff base (SB) is in the protein center. SB, D85 and D212 are the central cluster and E194 and E204 are the proton release cluster. (*) Lys159 and Glu161and the backbone connecting them is not resolved in 1C8R, and were added from the early M structure (1DZE).

The proton transfer pathway is divided in two by the Schiff base. On the extracellular side of the Schiff base, Arg 82, together with a number of polar groups, including Tyr 57, 83 and 185 and Ser 193, connects the central cluster and the proton release cluster. In contrast, the other half of the pathway, between the Schiff base and Asp 96 on the cytoplasmic side, is hydrophobic. No other ionizable residue is found within the ten Ångstrom long proposed proton transfer pathway.

.

The reaction cycle starts with retinal photo-isomerization from all-*trans* to 13-*cis* configuration [21,22] (Fig 1.3). The intermediates are designated bR (ground state), K, L, early and late M, N, and O. The sub-microsecond changes in K and L are localized to the retinal Schiff base (SB) itself and do not involve proton transfers. In the ground state the retinal is in its *trans* configuration, the SB is protonated and each acid in the central cluster ionized. In the proton release cluster one proton is bound. By M a proton has been transferred from the SB to Asp 85. And the proton on the release cluster is released into the extracellular space. The next step occurs in the cytoplasmic half-channel as the protonated Asp 96 reprotonates the SB forming the N intermediate. Following reisomerization of the retinal, Asp 96^{-} takes up a proton from the cytoplasmic side forming the O intermediate. Finally, proton transfer from Asp 85 to the extracellular proton release cluster reforms the ground state.

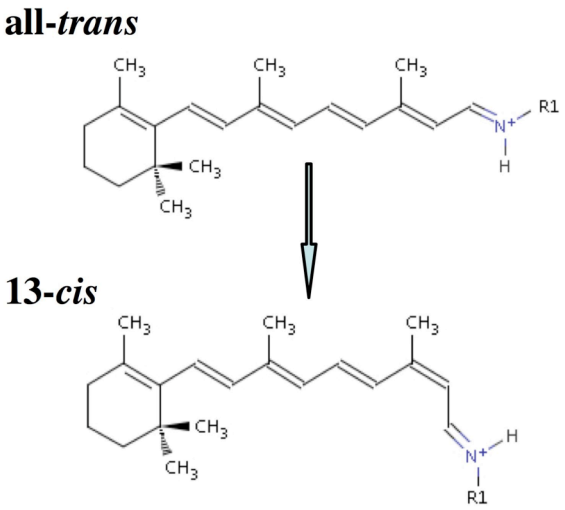

Figure 1.3 Schematic structure of a retinal in all-trans and 13-cis configurations

.

	Intra-cellular (High pH)	Asp 96	SB	Asp 85+212	Glu 194+204	Extra-cellular (Low pH)	SB*
bR		0	+1	-2	-1		T
M_{early}		0	**0**	**-1**	-1		C
M_{late}		0	0	-1	**-2**	***H⁺ gained***	C
N		-1	+1	-1	-2	*H⁺ gained*	C
O	***H⁺ lost***	**0**	+1	-1	-2	*H⁺ gained*	T
bR	*H⁺ lost*	0	+1	**-2**	**-1**	*H⁺ gained*	T

Table 1.1 Ionization states of functional residues in each bacteriorhodopsin intermediate

Bacteriorhodopsin reaction cycle showing five intermediates with 4 key protonation sites(21,22). Protons are bound from the intracellular and released to the extra-cellular sides of the membrane. In an energized membrane this moves protons from high to low pH regions. Charges on Asp 85 + 212 and on Glu 194 + 204 are summed. A group or cluster charge is in bold for the step where it has changed charge. SB gives the configuration of the retinal Schiff base (T: all-trans and C: 13-cis). L and K, the short lived intermediates that connect bR and M, have the same protonation states as bR and are not shown. The bR state is repeated at the top and bottom to emphasize the cyclic nature of the sequence.*

Bacteriorhodopsin has been the subject of many studies and has become a model system for study of proton transfer reactions. Still there are questions unanswered: How does proton transfer couple to structural changes in the protein? How does the protein ensure that each proton transfer is in the same direction, therefore to avoid reactions that will short-circuit the pumping? Without understanding the thermodynamic and kinetic properties of the reaction, these questions cannot be addressed.

.

One advantage of studying bacteriorhodopsin with computational techniques is the abundance of structural information. Using lipidic cubic phase crystallization (23), it has been possible to obtain a remarkable series of high-resolution bacteriorhodopsin structures trapped in various stages of the photocycle. Structures are available of bR (11,24,25), K (26,27), L (26,28), early (25,29) and late (24,30) M and O-like (31) intermediates. Comparison of these structures show only small differences, with an atomic RMS deviations of about 1 Å. The question is whether the changes required for proton transfer are trapped in these intermediate structures. In Chapter 2, MCCE calculations are carried out with the structures from the ground state, and early and late M states to show that the proton shifts are indeed stabilized by the trapped structural changes.

Studies in Chapter 2 also demonstrate that the enzyme reaction is coupled to the structural changes trapped in the different crystal structures. The analysis presented here shows that only when the essential changes are identified can the mechanism be understood. While there is a wealth of crystallography data for bacteriorhodopsin, most enzymes do not have crystal structures for more than one intermediate. Then, how can a reaction mechanism be derived from a static structure? One approach is to model the structural changes using information available from other studies. For example, in bacteriorhodopsin, the isomerization of retinal from all-trans to 13-cis was identified before crystallography data was available (32,33). The essential conformational change on Arg 82 was first predicted theoretically (34). In Chapter 3, proton transfers in bacteriorhodopsin are calculated with the structural models that are derived from one ground state structure to determine if these are sufficient to catalyze the movement of proton.

.

1.3 Cytochrome c oxidase

Cytochrome c oxidases, the most prevalent heme-copper oxidases, are the terminal electron acceptors in anaerobic organisms. These transmembrane proteins reduce dioxygen and convert the released chemical energy into an electrochemical gradient, either across the eukaryotic mitochondrial membrane or the bacterial cell membrane (16,35-37). In cytochrome c oxidases 4 electrons, provided by 4 cytochromes *c*, reduce dioxygen to water. The four protons used to make water are taken from the negative cytoplasmic side of the membrane, adding to the electrochemical gradient. Four additional protons are pumped across the membrane (38). Thus, each dioxygen molecule reduced by cytochrome c oxidase is coupled to the transfer of 8 charges across the membrane.

Each of the four electrons transferred to the active site to reduce dioxygen take the same pathway (Fig 1.4). In the first step, one electron is donated from cytochrome c to Cu_A, a di-copper center in subunit II which extends beyond the membrane on the proton release side of the protein. After receiving the electron, Cu_A reduces the 6-coordinate, low-spin, bis-His-Heme a, which is in the center of the membrane embedded subunit I. Heme a then passes this electron into the binuclear center (BNC), formed by a heme with a single His ligand (Heme a_3) and a Cu with three His ligands (Cu_B). The BNC and Heme a are deeply buried in the protein, approximately 20 Å from the proton input side and 15Å from the proton release side of the membrane.

.

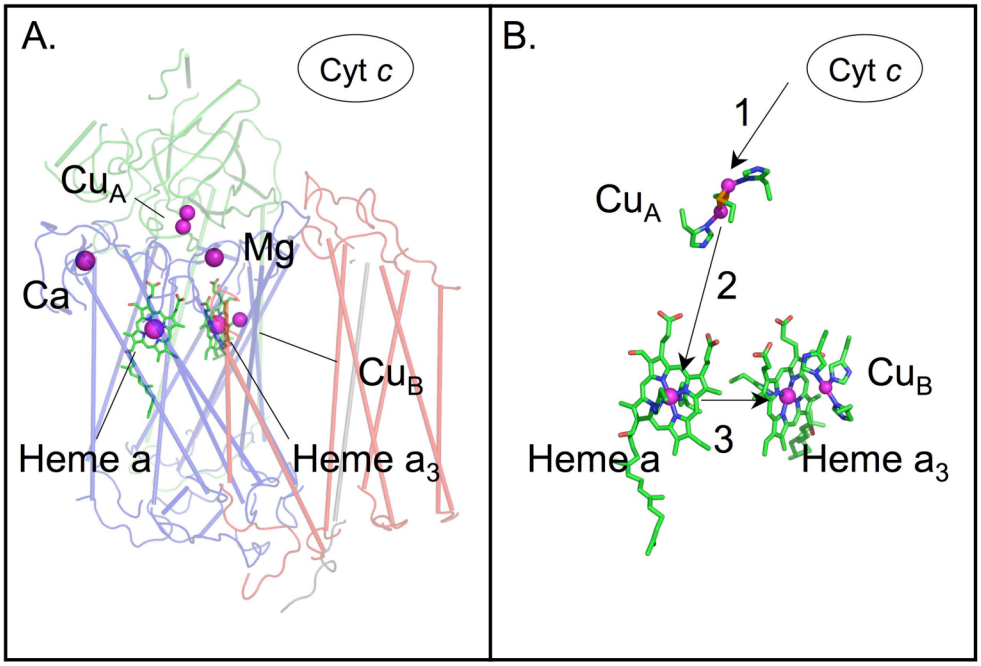

Figure 1.4 *Rb. sphaeroides* cytochrome c oxidase (PDB: 1M56)

A. Backbone trace of (blue) subunit I, (green) II, (red) III, and (grey) IV. Bound ions are shown as spheres. Cytochrome c in the solution phase is shown schematically above the oxidase structure. Heme a_3 and Cu_B make up the binuclear center (BNC). B. Electron transfer pathway goes from cytochrome c in solution to Cu_A, then to Heme a, and finally to the BNC.

Starting from a fully oxidized enzyme (O state), with Cu_A, Heme a and a_3 and Cu_B oxidized, the first electron is transferred to the BNC to form a singly reduced state (E state). Then the second electron transfers to the BNC forming the doubly reduced R state, with Cu_A and Heme a still oxidized, Heme a3 in the reduced ferrous form and Cu_B in the reduced cuprous state. One dioxygen molecule can now be bound to the BNC to form A state. No electron transfer occurs at R→A transition when dioxygen is reduced. The reduction of dioxygen is by four electrons in one step, to avoid forming any toxic oxygen radical. Other than the two electrons that are already donated to the BNC, two more electrons are needed. One of them comes from Heme a_3, which donates two electrons forming a ferryl (Fe^{IV}) species.

.

The other electron is borrowed from the protein, with Tyr I-288 the best candidate to donate one electron and a proton, forming the neutral radical species. After this super-oxidized oxidase is formed with 2 fewer electrons than the starting O state, it is then reduced by two more cytochromes c. The third electron donated by cytochrome c reduces Tyr I-288 forming the P state, while the fourth electron reduces Heme a_3 to the ferric state. The enzyme is now back to the fully oxidized state. The whole reaction cycle can be divided into two halves. The reductive phase, from O to R, occurs before O_2 binds to the BNC. Before dioxygen binds to the BNC, from O to R transition is the reductive phase. This half can be carried out anaerobically. The oxidative phase A to O transition is where the chemistry is carried out.

In order to produce water and pump protons, proton transfers need to be coupled to the O_2 reduction reaction. Two channels, D and K, have been identified for protons to travel from the input side of the protein into the BNC. The D-channel begins with Asp I–132 near the protein surface, and ends with Glu I–286 near the BNC (39-41). The K-channel begins with Glu II–101(42,43), passes Lys I–362 (44), and ends with Tyr I–288. Both channels are more than 20Å long and are assumed to be filled with water (45-47). It has been suggested that the single electron reduction of the fully oxidized oxidase is coupled to proton uptake via the K channel (40,48-50), and that the second reduction is coupled to proton uptake either via the D (40) or K channel (51).

The stoichiometry of proton transfer coupled to the electron transfers is still an open question despite numerous studies(34). It has been proposed that in the anaerobic reductive half of the reaction cycle, two protons are bound near the BNC (52-54). This requires two groups with in situ pK_as below the ambient pH in the oxidized enzyme; which allows them to be potential proton acceptors. Then when the cofactors are reduced, their pK_as would need to shift to be higher than the solution pH; so each can bind

.

a proton. If there is tight coupling between proton and electron transfers, then the reaction would be electroneutral, ensuring that the oxidized and reduced BNC complexes have the same net charge.

The standard view of electroneutrality is a reasonable assumption based on the cost of changing a charge on a site deeply buried in the membrane embedded protein. Electron transfer through oxidase changes the charge state of the redox cofactors. The favorable charge-dipole interaction of a charge with water, referred to as the solvation (reaction field or Born) energy, strongly stabilizes ionized states in water (55-58). If the charge is buried in the protein, then this favorable interaction is much smaller. Consequently, the loss of solvation energy always shifts reaction equilibria in a protein, favoring states with the a smaller net charge. The reduced Cu_A, Heme a and a_3 have a charge of zero, and there is a +1 charge on the reduced Cu_B. Cofactor oxidation in the protein, which increases the positive charge, is destabilized relative to the same redox reaction in aqueous solution by the solvation energy loss. Removal from water also shifts the pK_as of the surrounding acids and bases to favor their neutral forms. However, pairwise interactions with nearby anionic groups or dipoles can stabilize the buried charged residues and cofactors, compensating for the loss of solvation energy (59). The final in situ cofactor E_ms and residue pK_as reflect the relative contribution of these different terms. Significant perturbations, either stabilizing or destabilizing the charged forms, are common within proteins. For example, bis-His-hemes have E_ms ranging from –410 to +360 mV; exhibiting the ability of the protein to modify the chemical properties of buried groups. The E_m shifts for hemes with the same ligand have been shown to be caused by differences in the heme electrostatic environment within the protein (60-64).

There have been several proposed proton acceptors including: a hydroxide bound to Cu_B (52,65,66) or Heme a_3 (67), the Heme a_3 propionic acids (45,52,68-70), or one of the His ligands of Cu_B (71-73). In Chapter 5, MCCE is used to calculate proton uptake during the individual stages of anaerobic reduction of *Rb. sphaeroides* cytochrome c oxidase. The in situ pK_as and proton uptake were determined for all

.

residues. The stoichiometry of proton uptake for anaerobic reduction of the protein (51,74,75), and the E_m of Heme a with the BNC cofactors in different ionization states (54,76-78) were computed by MCCE and compared to published oxidase experiments. The main focus of the work is to determine, given the structure of the *Rb. sphaeroides* oxidase, the residues near the BNC whose equilibrium ionization states allow them to couple proton uptake to cofactor reduction, and to determine if there are two proton acceptors which can keep BNC reduction electroneutral.

One of the challenges that were encountered during this study was to determine the pK_a of the water bound to Heme a_3 in the reductive phase of the reaction. In a standard calculation, the calculated pK_a deviated from an experimentally determined pK_a by 6 ΔpK units. Chapter 6 focuses on a series of pK_a calculations with six different aquo-heme proteins, including oxidase, which reveals a breakdown in the classical continuum electrostatics analysis of pairwise interactions between a ferric water-heme and positively charged groups facing the heme plane. The interaction of the heme with these charges is compared in continuum electrostatics (CE) and density function theory (DFT) methods, showing that CE overestimates the interaction. With the appropriate corrections, MCCE calculations reproduce the experimental pK_as and E_ms, showing how the protein structures yield the observed in situ reaction energetics.

.

2. Methodology

2.1 Molecular Simulation

A protein is a complex heterogeneous system made of one or more chains of amino acids (polypeptides) folded together. Within a protein each amino acid can take different atomic positions; acidic and basic amino acids can have different protonation states; in addition, a protein can bind water and other cofactors, which also have different positions, protonation states and/or redox states. Therefore proteins are large systems with many degrees of freedom occupying complicated energy landscape. A full examination of the energy surface is not possible. This is demonstrated by the Levinthal paradox (79): considering a small protein of 100 residue, and each residue-residue joint has 3 configurations, the fastest motion, using 10^{-15} second to sample each microstate, takes 1.6×10^{25} years to sample all conformations, much longer than the age of the universe. Although Leventhal's paradox asks how nature can fold proteins within seconds after a polypeptide is synthesized, computer simulations face the same problem (80). Molecular simulation techniques somehow mimic nature to explore only a small set of configurations that can represent the energy surface. Molecular dynamics and Monte Carlo sampling are two simulation methods commonly used to analyze a whole protein.

In a molecular dynamics simulation, continuous configurations of a system are generated by integrating Newton's laws of motion. The result is a trajectory of position and speed of all atoms varying with time. Many thermodynamic quantities can be calculated with a molecular dynamics simulation, such as system energy, heat capacity, and temperature.

A Monte Carlo simulation generates configurations of a system by applying random changes. Techniques such as the Metropolis algorithm enable the simulation to sample the energy landscape based on the energy. The result is a population distribution of different configurations that reproduce the Boltzmann distribution.

.

The most obvious difference between molecular dynamics and Monte Carlo is that molecular dynamics provides time-dependent information. While in Monte Carlo, each change is random, without any information on how much time it would take in a real system. The advantage for molecular dynamics is that the kinetic properties can be calculated. However in cases where regions of energy surface to be sampled are separated by large barriers, Monte Carlo technique has the advantage of reaching equilibrium more efficiently with a low cost.

In the work reported here, different protein configurations are explored to determine the protonation states of all ionizable residues. Ionization state change is a chemical reaction, which cannot be modeled by a classic molecular dynamics simulation. Hybrid programs joining molecular dynamics with a quantum mechanics calculation at a site of interest can calculate proton and electron affinity using the QM/MM technique (81-86). However it is a challenge to connect the quantum mechanics (QM) and molecular mechanics (MM) regions of the simulation to estimate the electrostatic impact of the rest of protein on the protonation site (87). In addition, the protonation states of the MM region still cannot be sampled in these calculations.

The Monte Carlo technique can sample changes of ionization states in a grand canonical ensemble, where chemical potential μ is fixed and the number of particles is a variable. The chemical potential variable for changing a proton is pH, and for changing an electron is E_h.

2.2 Electrostatics calculation

Protein is formed of atoms with different electronegativity, which yields an unsymmetrical electron distribution. As a result, an electrostatic field is generated, driving the motions of atoms and electrons. Although the charge distribution is raised by nuclei and electron density throughout the space, a simpler model is used to describe the charge distribution in molecular simulations. In this model,

.

fractional charges are assigned to atoms to reproduce the electrostatic field, so called partial charge model.

The electrostatic potential given by a charge q1, and its interaction energy with another charge in the space q2 can be described by Coulomb's law in a vacuum or medium of uniform ε:

$$V(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \bullet \frac{q_1}{|\mathbf{r} - \mathbf{r_1}|} \tag{2.2.1}$$

$$E = \frac{1}{4\pi\varepsilon_0} \bullet \frac{q_1 q_2}{|\mathbf{r_2} - \mathbf{r_1}|} \tag{2.2.2}$$

However protein is surrounded by water instead of a vacuum. Solvent molecules can be polarized, such as water molecules, and they are highly mobile, so a solvent such as water has a large polarizability. When a solute is moved into a solvent from a vacuum, solvent molecules re-equilibrate around the solute. The energy difference moving a solute from vacuum to a solvent is called solvation energy. Explicitly calculating the energy requires modeling a large number of solvent molecules and their dynamic properties with and without the solute. In addition, a solute also has polarizable electron density and fluctuating nuclei, which adds to the cost of the calculation. It is a challenge to explicitly calculate the solvation effect when studying a large protein.

To calculate the solvation energy with less cost, both solute and solvent are treated as continuum material with different dielectric constants. Two models are commonly used in molecular simulations, the generalized Born model and the continuum electrostatics (Poisson-Boltzmann) model.

In the Born model, the self-energy of a spherical object with a radius r (Born radius), having charge q uniformly distributed on the surface, is $\frac{1}{4\pi\varepsilon_0} \bullet \frac{q^2}{2\varepsilon r}$ in a material with dielectric of ε. The

.

energy change moving this object from a vacuum to the medium with dielectric constant ε is $-\frac{1}{4\pi\varepsilon_0}\bullet\frac{q^2}{2r}\left(1-\frac{1}{\varepsilon}\right)$.

In a protein, the protein-solvent boundary is not spherical, but has an irregular shape. The analytical solution of Born model can no longer be used. The generalized Born model was developed to use the same type of formulism to calculate the solvation energy. In this model, effective Born radius is used to estimate the effect of the size and geometry of the protein and the burial of the charged atom. Using the generalized Born model, the solvation energy can be calculated rapidly so is commonly used in molecular dynamics simulations.

The continuum electrostatics model assigns different dielectric constants to the solvent and solute, and then solves the Poisson-Boltzmann equation to calculate the electrostatic field and the solvation energy. The electrostatic field in a position dependent dielectric obeys the Poisson equation,

$$\nabla\bullet\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})=-\rho(\mathbf{r}) \tag{2.2.3}$$

In order to model mobile ions in the solvent, the Poisson equation is modified to the Poisson-Boltzmann equation. The mobile ions in the solvent are modeled as a charge density distributed in a Boltzmann distribution under the electric field, which introduces an extra term to the Poisson equation:

$$\nabla\bullet\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})-k'\sinh[\phi(\mathbf{r})]=-\rho(\mathbf{r}) \tag{2.2.4}$$

where $k'=\sqrt{\frac{8\pi N_A e^2 I}{1000 k_B T}}$ and e is the electronic charge, I is the ionic strength of the solution and N_A is Avogadro's number. Delphi (88) is one the software packages which can numerically solve the Poisson-Boltzmann equation.

.

2.3 Multi-Conformation Continuum Electrostatics

Protein flexibility has been recognized as being important for determine the protein electrostatic environment (89), ligand binding (90,91) and other protein functions (92). The static crystal structure does not provide a dynamic picture of an active protein. To study protein function, it is important to have a tool that can start with the initial crystal structure, and then model structural changes explicitly.

Electrostatic effects are one source of the correlation between protein structure and function (93-97). Therefore it is important to calculate electrostatic energies accurately. The Poisson or Poisson-Boltzmann (PB) equation is considered the most accurate method to calculate these energies (98,99).

Crystal structures from the Protein Data Bank (100,101) usually do not have sufficient resolution to resolve the position of protons. Changes of ionization states, proton and electron transfers, play an important role in protein functions (102). A computational tool needs to be able to determine the ionization states and their changes coupled to protein reaction.

Multi-Conformation Continuum Electrostatics (MCCE) is a hybrid method developed to meet all these requirements. MCCE includes structure flexibility by adding rotamers to the structure (103). The entire sidechain rotamer phase space is searched (104). MCCE uses Delphi program to numerically solve Poisson-Boltzmann equation (98,99,105). And finally by sampling all the configurations added to the structure, a Boltzmann distribution of residue ionization, positions, and water or ligand binding are determined.

In the standard MCCE scheme, a protein is broken down to separate residues (amino acids and cofactors). Each residue is given different positions and ionization states, which are called conformers. MCCE treats residue motions with a number of pre-selected positions to evenly sample the phase space. A protein microstate is formed by each residue occupying one conformer. The energy of microstate n (ΔG^n) is the sum of the electrostatic and non-electrostatic energies (63) defined by:

.

$$\Delta G^n = \sum_{i=1}^{M} \delta_{x,i} \{ [2.3 m_i k_b T (pH - pK_{sol,i}) + n_i F (Eh - Em_{sol,i})] + (\Delta\Delta G_{rxn,i} + \Delta G_{pol,i}) \} + \sum_{i=1}^{M} \delta_{x,i} \sum_{j=i+1}^{M} \delta_{x,j} [\Delta G_{ij}] \quad (2.3.1)$$

where $k_b T$ is 0.59 kcal/mol (25.8 meV), M is the number of conformers, $\delta_{x,i}$ is 1 for conformers that are present in the state and 0 for all others. m_i is 1 for bases, –1 for acids, and 0 for polar groups and waters. n_i is the number of electrons gained or lost compared to the ground state conformer. For example, if an oxidized conformer is defined as the ground state, it has $n_i = 0$ and a reduced conformer has $n_i = 1$. F is the Faraday constant. $pK_{sol,i}$ is the pK_a of the ith group in solution. $E_{m,sol,i}$ is the midpoint potential of the ith cofactor in solution. $\Delta\Delta G_{rxn,i}$ is the difference between the conformer reaction field energy in solution and in the protein (desolvation energy). $\Delta G_{pol,i}$ is the pair-wise electrostatic and non-electrostatic interaction of the conformer with the backbone, and with side chains that have no conformational degrees of freedom. The torsion energy for each conformer is included here. ΔG_{ij} is the electrostatic and Lennard-Jones pair-wise interaction between each pair of conformers in the microstate. The limits on the summation of the inter-conformer terms ensure that each interaction is counted once.

Each energy term is pre-calculated to make an energy lookup table, once all the conformer choices are made. $pK_{sol,i}$ and $E_{m,sol,i}$ are defined based on residue type and can be measured experimentally. The electrostatic interactions ($\Delta\Delta G_{rxn,i}$, $\Delta G_{pol,i}$, and the electrostatic part of ΔG_{ij}) are calculated by solving the Poisson-Boltzmann equation using multiple DelPhi (88) runs integrated into the MCCE program. This model for electrostatics calculations treats solute as a continuum material with a low dielectric constant, and solution as a continuum material with a high dielectric constant. Charge distribution inside the protein is described by atomic charges of each residue. The mobile ions in the

.

solution redistribute in response to the electric field raised by the protein and the ions. The ion distribution can be described by a Boltzmann distribution of charge density under the given potential. The non-electrostatic interactions include Lennard-Jones interaction and torsion energy. All atom-atom Lennard-Jones interactions are calculated. Lennard-Jones interactions for those that are directly bonded (1-2 interaction), or connected by two bonds (1-3 interaction) are set to zero, and for those that are connected by three bonds (1-4 interaction) are reduced. Torsion energies are defined by four atoms connected by sequential bonds.

A set of parameters has been defined for a standard MCCE calculation. All the amino acid $pK_{a,sol}$ are taken from studies of peptides (106,107). The low dielectric constant for the solute is 4, the high dielectric constant for the solution is 80. The average ionic strength in the solution is 0.15mM. PARSE charges (108) are used for the atomic charges of all amino acids. PARSE radii (108), a solvent accessibility based parameter, are used to draw the boundary between low dielectric solute and high dielectric solution. The non-electrostatics interactions are calculated with AMBER force fields. In the Monte Carlo sampling all Lennard-Jones interactions are divided by 4. This rescaling is needed because the Lennard-Jones parameters are optimized with electrostatic interactions calculated with $\varepsilon=1$, while in MCCE $\varepsilon=4$ is used. A smaller Lennard-Jones repulsion is needed with $\varepsilon=4$ to ensure that hydrogen bonds have the correct distance dependence.

The latest version of Multi-Conformation Continuum Electrostatics (MCCE2) (104) is compiled recently with improved rotatmer making and optimization. The details of the developed subroutines are described in the appendix section.

.

2.4 Material and Methods for calculations of bacteriorhodopsin

Atomic coordinates. Bacteriorhodopsin ground state structures from the protein data bank (101) 1C8R (24), 1C3W (11), 1KG9 (29), 1M0M-1 (model 1 in 1M0M) (25) were analyzed. The early M structures 1DZE (Takeda et. al.), 1KG8 (29), 1M0M-2 (model 2 in 1M0M)(25), illuminated at ≈220 K until the purple to yellow transition has taken place then frozen to 100K, were used. The late M structures 1C8S (24) and 1F4Z (30) were analyzed. These crystals were frozen to 100K and illuminated while the cyrosolvent stream was blocked for several seconds. 1C8S is a D96N mutant and 1F4Z an E204Q mutant, each of which slows the transition from M to N. There are only small differences between the atomic positions in the different coordinate files. The RMSD amongst the three structures 1C8R (bR), 1DZE (early M) and 1C8S (late M) are 0.5~0.7Å for backbone and ~1 Å for backbone and side chains. Except for 1M0M-2 all M structures have the SB nitrogen pointing towards the cytoplasm and Arg 82 pointing towards the extra-cellular space. Mutated residues in the structure files were replaced by the wild-type residues in the analysis. The 1M0M-1 structure has a very non-standard conformation for Asp 212 which was relaxed in Swiss-PDB Viewer (109). No other minimization was carried out.

A full length bacteriorhodopsin has 248 residues. No structure is complete. Missing residues in the N and C termini and the E-F loop could change the results by exposing the protein interior to solvent or by removing long-range electrostatic interactions. Comparison of the loss in reaction field (desolvation) energy in the smaller 1C8S and the other proteins shows the deleted residues do not significantly change the solvent exposure of central or proton release cluster residues. The loss of pair-wise interactions with missing residues also creates only minor errors.

The retinal in bacteriorhodopsin is covalently attached to the side chain nitrogen of Lys 216, forming a Schiff base (SB) which can be neutral or protonated. In bR the retinal is in the all-trans

.

configuration while in M it is 13-cis. The same atomic charges(110) were used for both isomers. The SB pK_{sol} is set to 7 (111).

Waters with more than 10% of their surface exposed without a membrane slab are always deleted from the structure file in MCCE (112). For calculations with continuum waters in cavities all waters were removed. Added side chain rotomers can fill internal cavities. Calculations with continuum water were run twice. In the second run all buried conformers that were not selected by Monte Carlo sampling were deleted. This provides a more reproducible water cavity while maintaining important side-chain positions.

Calculations with explicit waters were carried out with 1C8R, 1DZE and 1C8S. All buried crystallographic waters were included. Extra water oxygens were added at every fifth IPECE lattice point (2.75Å) to fill cavities. For the central and exit cluster extra water oxygens were added to make good hydrogen bonds to the protein. Standard MCCE algorithms added protons to the oxygens so waters can donate or accept hydrogen bonds to neighboring residues (103). Each water oxygen gets ≈10 conformers differing in their proton position. Each water also has a conformer with no interactions with the protein representing movement into the bulk. The atomic charge is -0.8 on oxygen and 0.4 on each hydrogen. The reaction field energy of an individual water in bulk water is –1.30 ΔpK_a unit.

Reduced model. The 400-500 conformers in each structure generate an extremely large number of possible microstates. This makes it difficult to identify the factors that control ionization changes. A reduced model is used to simplify the complexity. Here analysis focuses on a single cluster composed of several residues where strong interactions yield interdependent ionization states. Conformational changes of non-cluster residues, which are weakly coupled to cluster residue ionization states, are ignored. The non-cluster residues interact with the cluster via a mean-field approximation given their occupancy in a full MCCE calculation. For the cluster, only the most populated ionized and neutral

.

conformers in MCCE are considered. Therefore for n residues in the cluster, there are 2^n ionization states. So for SB, Asp85 and 212 in the central cluster there are 8 microstates and for Glu194 and 204 in the proton release cluster there are 4 microstates. The primary advantage of the reduced model is that the number of microstates is now enumerable so all state energies can be compared. In the reduced model, the M conformers are split into m residues in the cluster and M' surrounding residues. The pair-wise energy in a given cluster ionization microstate is:

$$\left\{\sum_{i=1}^{m}\delta_n(i)\sum_{j=i+1}^{m}\delta_n(j)\left[\Delta G_{ij}\right]\right\} + \left\{\sum_{i=1}^{m}\delta_n(i)\left\{\sum_{j=1}^{M'}\rho(j)\left[\Delta G_{ij}\right]\right\}\right\} + \left\{\sum_{i=1}^{M'}\rho(i)\sum_{j=i+1}^{M'}\rho(j)\left[\Delta G_{ij}\right]\right\} \qquad (2.4.1)$$

When calculating the full microstate energy in the reduced model equation 2 replaces the third line in equation 1 which contains the conformer-conformer pairwise interactions in the full MCCE energy function. Here $\rho(j)$ is the Boltzmann probability of non-cluster conformer j in the full MCCE calculations. The first term contains the interaction amongst cluster conformers, calculated explicitly; the second term, $\sum_{j=1}^{M'}\rho(j)\left[\Delta G_{ij}\right]$, has the mean-field interactions between cluster conformer *i* and all non-cluster conformers; the third term has the interactions between non-cluster conformers, which is the same for all cluster microstates.

In addition to determining the total microstate energy the energy of each residue in each microstate can be determined. Without intra-cluster interactions, the protein shifts the energy of a given conformer i by $\Delta\Delta G_{rxn,i} + \Delta G_{pol,i} + \sum_{j=1}^{M'}\rho(j)\left[\Delta G_{ij}\right]$, the first two terms are same as in full MCCE and the last term is the mean-field energy from non-cluster residues. The energy difference between ionized and neutral conformer of residue k is:

.

$$\Delta G'_k = \left\{ \left(\Delta\Delta G_{rxn,k(I)} + \Delta G_{pol,k(I)} + \sum_{j=1}^{M'} \rho(j) \Delta G_{jk(I)} \right) - \left(\Delta\Delta G_{rxn,k(N)} + \Delta G_{pol,k(I)} + \sum_{j=1}^{M'} \rho(j) \Delta G_{jk(N)} \right) \right\} \quad (2.4.2)$$

So without intra-cluster interaction, the pH where ionized and neutral conformers of residue k are of equal energy is:

$$pK'_k = pK_{sol,k} - \gamma_k \Delta G'_k \quad (2.4.3)$$

where pK_{sol} is the pK_a of the ionizable group in the solution, γ is 1 for bases, -1 for acids. Now the total energy of a given microstate in the reduced model is:

$$\Delta G_n = \sum_{i=1}^{m} \delta_n(i) \left[\gamma_i \left(pK'_i - pH \right) \right] + \sum_{i=1}^{m} \delta_n(i) \sum_{j=i+1}^{m} \delta_n(j) \left[\Delta G_{ij} \right] \quad (2.4.4)$$

where γ is 0 for all neutral conformers, 1 for ionized bases and -1 for ionized acids.

2.5 Material and Methods for calculations of cytochrome c oxidase

Subunits I-IV of the first model of *Rb. sphaeroides* cytochrome c oxidase structure in the PDB file 1M56 (113) are analyzed with MCCE. Water molecules are removed and the internal cavities treated as a high dielectric space in the electrostatics calculation. A 32 Å slab of neutral atoms is added to provide a low dielectric membrane, with the slab position optimized to bury the fewest surface ionizable residues using the program IPECE (114).

Parameters for E_m and pK_a calculations of the cofactors. Cu_A is a di–copper complex with eight sidechain and backbone groups serving as ligands. The copper atoms and ligands are treated as a complex, and the charge distribution is obtained with Gaussian98 (115). The B3LYP method (116) with LANL2DZ basis set (117) was used, and the CHELPG (118) algorithm used to fit atomic charges. This

.

charge set is able to recover measured E_ms of the designed di-Cu-azurin and of the oxidase soluble domain from different species (manuscript in preparation).

Both Heme a_3 and Cu_B have one open coordinate position to bind water or hydroxyl. Conformers are built for water–Heme a_3, hydroxyl–Heme a_3, water–Cu_B, and hydroxyl–Cu_B. On Heme a_3, the Fe–O bond is oriented perpendicular to the porphyrin plane, and the bond length is 1.95Å with square bipyramidal geometry (119). The oxygen position for Cu_B is determined by DFT optimization calculations using the B3LYP method (116) and LANL2DZ basis set (117) for an isolated cupric hydroxyl-3His-Cu complex. The C_α and C_β atoms of all His ligands and the Cu are fixed in their crystal positions during optimization. This yields a distorted tetrahedral Cu geometry. The resultant O-Cu-N_ε(His I-333) bond angle is 136º and the Cu-O bond length is 1.9Å. Hydrogens are located with tetrahedral oxygen geometry and bond lengths of 0.96Å. On each cofactor, conformers of water and of hydroxyl are created by rotating the protons by 30º around the Fe–O or Cu–O bond, yielding 12 water and 12 hydroxyl positions on each cofactor. Thus, there are 48 different conformers for Heme a_3 and for Cu_B given their two ionization states each with 24 possible water or hydroxyl positions. Cu_B has an additional set of conformers described below, which include the possible ionization change of a His ligand.

The approach used here to calculate the E_ms of cofactors and pK_as of the water ligands treats the metal and their ligands as a complex. The assigned $E_{m,sol}$ and $pK_{a,sol}$ is taken from the measurements of a model system for the complex in solution. The reaction field energy is also calculated treating the whole complex. Any non-bonded pairwise interactions amongst the metal and its ligands are ignored since these interactions are included in the $E_{m,sol}$ and $pK_{a,sol}$. Treating cofactors and ligands as a complex thus avoids the errors inherent in using classical electrostatics to calculate the bonding energies joining them together. However this approach requires prior knowledge of $E_{m,sol}$ and $pK_{a,sol}$ for the desired complex.

.

In this study, an $E_{m,sol}$ of –120 is assigned to the a-type bis-His heme found in the Heme a site (120,121), 80 mV higher than for a c–type bis-His heme microperoxidase 8 (MP8) (122,123). An a-type heme has an extra electron-withdrawing formyl group that favors reduction (64). The assigned $E_{m,sol}$ for Heme a_3 is based on the measurements of a heme c MP8 with a single His ligand. The $E_{m,sol}$ for the His-H_2O MP8 is –140mV (124,125) and –204mV for the His-OH^- MP8 (124). Since it is assumed the a-type heme E_ms are 80 mV more positive than the c-type MP8 with the same ligands, –60mV and –124mV are used for $E_{m,sol}$ for His-H_2O and His-OH^- Heme a_3. For calculating the pK_a of water ligands, a $pK_{a,sol}$ of 9.6 is assigned to the oxidized ferric His-aquo-heme and 10.9 for the reduced ferrous His-aquo-heme, which has been measured with a c-type MP8 heme (126,127). It is assumed that water pK_as are the same in c-type and a-type hemes because the water to be protonated is not in the porphyrin plane so should not be influenced by the additional electron withdrawing group on the porphyrin edge. There are fewer measurements that can be used to model the water ligand pK_a in the Cu_B complex. A $pK_{a,sol}$ of 9.4 is used for aquo-Cu_B(II). This is taken from measurements of a tripodal ligand tris[2-(methylamino) ethyl] amine complex (128,129), which is similar to the Cu_B complex. A measured $pK_{a,sol}$ of aquo–Cu_B(I) has not been found in the literature. It is likely to be higher than that of Cu_B(II) due to the smaller positive charge on the metal. A conservative $pK_{a,sol}$ of 9.4 will be assigned to aquo–Cu_B(I) and the effects of this choice discussed below.

A metal centered charge set is used for Heme a, Heme a_3 and Cu_B. Heme a and a_3 have a +2 or +3 charge placed on Fe and –0.5 on each N atom of the porphyrin. Cu_B has a +1 or +2 charge on Cu. PARSE charges are used for the neutral His ligands. TIP3 charges are used for water ligands. CHELPG charges for the hydroxide and a His^- are calculated in isolation with the B3LYP method using the 6–31G* basis set. The metal centered charge distribution differs from that determined by DFT calculations for the complex where some of the metal positive charge is shifted to the ligands. A +0.3 charge is

.

placed on the formyl group C and -0.3 on the of the electron-withdrawing formyl group of the a-type hemes. This simple metal centered charge set has been used successfully in heme benchmark calculations for bis-His and His-aquo hemes (61,63,119).

It has been proposed that His I-334, a ligand to Cu_B, forms an anionic imidazolate during the reaction cycle (72,73). Two approaches are used to calculate the pK_a of this His ligand. The first is the same as used for the pK_a calculations of the BNC water ligands. Cu_B and all its His and aquo-ligands are treated as one complex. There is no experimental $pK_{a,sol}$ available for such a complex with a deprotonated imidazolate. However a $pK_{a,sol}$ of 9.0 has been calculated with DFT for a isolated Cu_B complex with water and 3 His ligands (73). A conformer with ionization state H_2O–His^-–Cu_B(II) is added to compete with H_2O–His^0–Cu_B(II) and OH^-–His^0–Cu_B(II) ionization states to determine the in situ His I-334 pK_a. The DFT calculations did not analyze ionization energies with both His^- and OH^- both bound on CuB(II) or any pK_as with a reduced Cu_B so these states cannot be addressed treating Cu_B as a complex with its ligands. An alternative approach allows estimation of the His pK_a in the presence of OH^-, the OH^- in the presence of His^- and the pK_as with reduced Cu_B. Here His is treated as an isolated group separate from Cu_B with its bound water and the two other His ligands. The $pK_{a,sol}$ of the free imidazolate of 14.4 is used. The reaction field energy of the His and the smaller Cu_B complex are each calculated. Explicit non-bonded electrostatic interactions between His or His^- and the different redox and protonation states of the smaller Cu_B complex are included. This approach has been shown to work for some heme ligands (61), but with a metal centered charge, continuum electrostatics often overestimates the favorable interactions between a metal and an anionic ligand.

Based on a study of aquo–heme pK_as and E_ms in various proteins (119), the Poisson-Boltzmann continuum electrostatics calculations are found to overestimate the interactions of ferric Heme a_3 with Cu_B. That study compared the energy of a charge near a ferric aquo-heme in vacuum calculated with

.

DFT or with Coulomb's Law. For the neutral, ferric His-hydroxyl-heme, Coulomb's law reproduces the change in the heme energy found in the DFT calculations within 10%. The correlation between the Coulomb interaction and the DFT energies has a slope of one, and is independent of the position of the external charge. Thus, the DelPhi continuum electrostatics calculations in MCCE should provide the correct interaction energies. However, there are position dependent systematic errors in interactions with the cationic, His-water-heme complex. While interactions with negative charges, and with charges in the heme plane, such as the propionic acids and with distant charges, such as Heme a, are well represented by Coulomb's law, when a charge is near the face of the heme the interactions are overestimated. Interactions with groups close to the heme also need correction in the benchmark calculations of the aquo-heme pK_as in sperm whale and *aplysia* myoglobin, hemoglobin I, heme oxygenase 1, and horseradish peroxidase (119).

The interaction between Heme a_3 and Cu_B is investigated with DFT. Different Heme a_3 and Cu_B redox states, ligand protonation states, hydrogen positions created in MCCE are tested. The correlation between DFT and Coulomb's interaction cannot be described by a simple scaling factor. But, these results showed the ranking of the energy states agrees in the two calculations for different hydrogen conformations. Thus, MCCE calculations are first carried out using uncorrected Poisson-Boltzmann (PB) interactions to find the equilibrium hydrogen positions in each BNC redox and protonation state. The DFT and Coulomb's law interactions are then calculated with the equilibrium conformations. The scaling factor required to bring the Coulomb's law interactions into agreement with the DFT calculations is applied to PB interactions and the scaled values are used in the pK_a calculations. Conformers that are not occupied in the calculations with uncorrected PB interactions are omitted in the pK_a calculations.

.

2.6 Material and Methods for calculations of aquo-heme pK_a

Structures of sperm whale myoglobin (1A6G, 1A6K, 1A6M, 1A6N, 1HJT, 1JP6 and 1JP9), *Aplysia* myoglobin (1MBA, 5MBA), monomeric clam hemoglobin I (1B0B, 1EBT, 1FLP and1MOH), rat heme oxygenase 1 (1DVE, 1IVE, 1IX4, 1ULX and 1VGI), horseradish peroxidase (1ATJ, 1H55, 1H57, 1H58, 1H5A, 1H5C, 1H5D, 1H5E, 6ATJ and 7ATJ) and *Rb. sphaeroides* cytochrome c oxidase (1M56) were obtained from the Protein Data Bank (101). All crystal waters are deleted and protein cavities filled with continuum water. Multi-Conformation Continuum Electrostatics (MCCE) samples both residue ionization state and rotamer position as a function of pH and E_h (103,112). The cytochrome c oxidase calculations use the model generated for analysis of proton uptake coupled to anaerobic reduction in MCCE2.0 (119). In this larger protein a rotation angle of 120° is used to make sidechain rotamers, reducing the number of conformers. All other calculations use MCCE 2.2, which adds local optimization of side chain positions, pruning to delete energetically indistinguishable conformers and corrections of the CE energy terms for the errors in the dielectric boundary due to extra surface conformations (see www.sci.ccny.cuny.edu/~mcce). Preselected sidechain rotamers are generated in 60° increments around each rotatable bond for all amino acids. In each protein, a water or hydroxyl is added to the open heme coordinate position. The aquo-oxygen is placed on the heme iron with square bipyramidal geometry and a 1.95Å Fe–O bond length. Hydrogens are added in a tetrahedral geometry with a 0.96Å H–O bond length. Then additional hydroxyl and water conformations are generated in 30º increments around the Fe–O bond, creating 12 conformers for both water and hydroxyl.

MCCE calculates the shift in pK_a or E_m when a group is moved from aqueous solution to the protein. The solution pK_a ($pK_{a,sol}$) and E_m ($E_{m,sol}$) is obtained from measurements in water. $pK_{a,sol}$ for amino acids are taken from studies of peptides (106,107). The heme propionic acids are treated as described previously with a $pK_{a,sol}$ of 4.9 (63). The five-coordinate hemes are assumed to bind a water or

.

hydroxide as the sixth ligand. Heme and its His and water or hydroxide ligands are treated as a complex. A $pK_{a,sol}$ of 9.6 for the hydroxyl in the ferric His-aquo-heme has been measured in microperoxidase 8 (MP8) which has a heme c (126,127). The $E_{m,sol}$ for the His-water MP8 is –140 (124,125) and –204 mV (124) for the His-hydroxyl MP8. It is assumed that with the same ligands the E_m of heme a is 80 mV more positive than heme c (64,120,121) and that the water $pK_{a,sol}$ are the same for hemes a and c. No explicit pairwise interactions of water or hydroxide ligands, with the His ligand or heme to which they are bound are considered since these interactions are included in $E_{m,sol}$ and $pK_{a,sol.}$ The reaction field energy in both solution and the protein were calculated for the aquo-heme complexes. A systematic shift of 0.74ΔpK unit (1 kcal/mol) is applied to the energy of ferric hydroxyl-heme and –0.5ΔpK unit (–0.67 kcal/mol) to ferrous water-heme. This raises all ferric heme pK_as by 0.74 pH units and raises the aquo-heme E_ms by 30mV.

MCCE uses continuum electrostatics to calculate the interaction between the aquo–heme and charges and dipoles in the protein. Selected interactions with the ferric heme porphyrin together with its imidazole and water or hydroxyl ligand were also calculated using Gaussian 98 (115). The heme and imidazole are taken from the crystal structure of sperm whale myoglobin (1A6K), with water or hydroxyl added in GaussView (115). Each structure is then optimized using the B3LYP method (116) with the LANL2DZ basis set (117). Hydroxyl-hemes are expected to be low-spin (67,130) while water-hemes are generally high-spin (67,131). The water-heme is given a net charge +1 and spin 5/2, while the hydroxyl-heme has a net charge 0 and spin 1/2. The change in system energy is then calculated with an external charge of +1 or –1 placed next to the aquo-heme complex in positions shown in Fig 2.1. All atoms are fixed in the structure optimized before the charge is added.

Nine external charges are placed on a plane parallel to the porphyrin plane located 4, 6, 8, 10 or 12Å from the heme (Fig 2.1). Points that overlap the water or imidazole ligand are removed. The

.

interactions are also calculated with Coulomb's law with the metal centered charges used in the MCCE calculations with ε=1. For the ferric hydroxyl-heme, Coulomb's law reproduces the DFT interactions with an error of <10% and a slope of 0.94 with the exception of points within 4 Å of the heme plane (Fig 2.2). However, for the water-heme, the Coulomb's law interactions with the points above the heme plane are significantly larger than those obtained with DFT. In contrast, for charges on the heme edge (Fig 2.1) the Coulomb's law interactions reproduce the DFT interactions with <5% error and a slope of 1.0 for both water and hydroxyl ligated hemes (Fig 2.2). The source of the error in the CE calculation is not known. The correction could be needed because there is significant charge at positions that are not well described by the atom-centered charges used here. Alternatively, there could be a different degree of polarization of the ferric water- and hydroxyl- hemes by the external charge. Further studies are in progress to develop a water-heme charge distribution, perhaps including charges out of the heme plane, which can better reproduce the DFT interaction with Coulomb's law. Preliminary calculations show that in general there is no systematic error in the CE calculations of the interactions of the standard amino acids with external charges.

The ferric water-heme interaction of a simple external charge at the position of Arg 38 in peroxidase or Cu_B in oxidase calculated with DFT is only 60% of that found with Coulomb's law. The interaction with a charge at the position of Arg 99 in hemoglobin I is 65% as large. The influence of the charge distribution for the external group on the correction was examined by using a PARSE Arg charge distribution, rather than a single +1 charge. The DFT interaction is now 63% of that found with Coulomb's law, very similar to that found with a unitary charge model. The influence of the charge distribution of Cu_B is more complex and cannot be described by a simple scaling factor. The details are discussed in ref. (132).

.

Corrections are applied to interactions of the water-heme with Arg 99 in hemoglobin, Arg 38 in peroxidase and Cu_B in oxidase. However, pairwise electrostatic interactions in MCCE are calculated with the Poisson-Boltzmann equation (PB), a CE analysis which takes into account the different dielectric constants of the protein and water (133,134). The scaling factor derived for particular heme-residue geometry from the comparison of Coulomb's law and DFT interactions in vacuum (Fig 2.2) is used to scale the individual PB interactions in MCCE. Corrections used for Cu_B can be found in ref. (132).

Interactions with residues that are on the heme plane or far away from the heme are shown to need no correction. Residues that have significant interactions with the aquo-heme which are not corrected because of their position include: all the heme propionates; Arg 45 which is hydrogen bonded to a propionic acid in myoglobin; Lys 179 which is hydrogen bonded to a propionate in hemoglobin 1; Arg 136 and Asp 140 which form an ion pair on the porphyrin edge furthest from the propionates in hemoglobin1 (see Chapter 7).

.

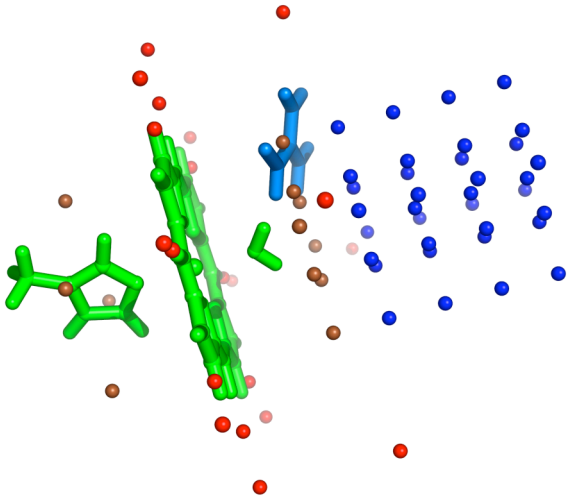

Figure 2.1 Position of external charges for comparison of CE and DFT interactions

Blue points are near the heme face; red in the plane of the heme and near the heme edge. Those within 4Å of the heme plane are brown. His-aquo heme is green. The position of Arg 38 in horseradish peroxidase is shown in blue.

.

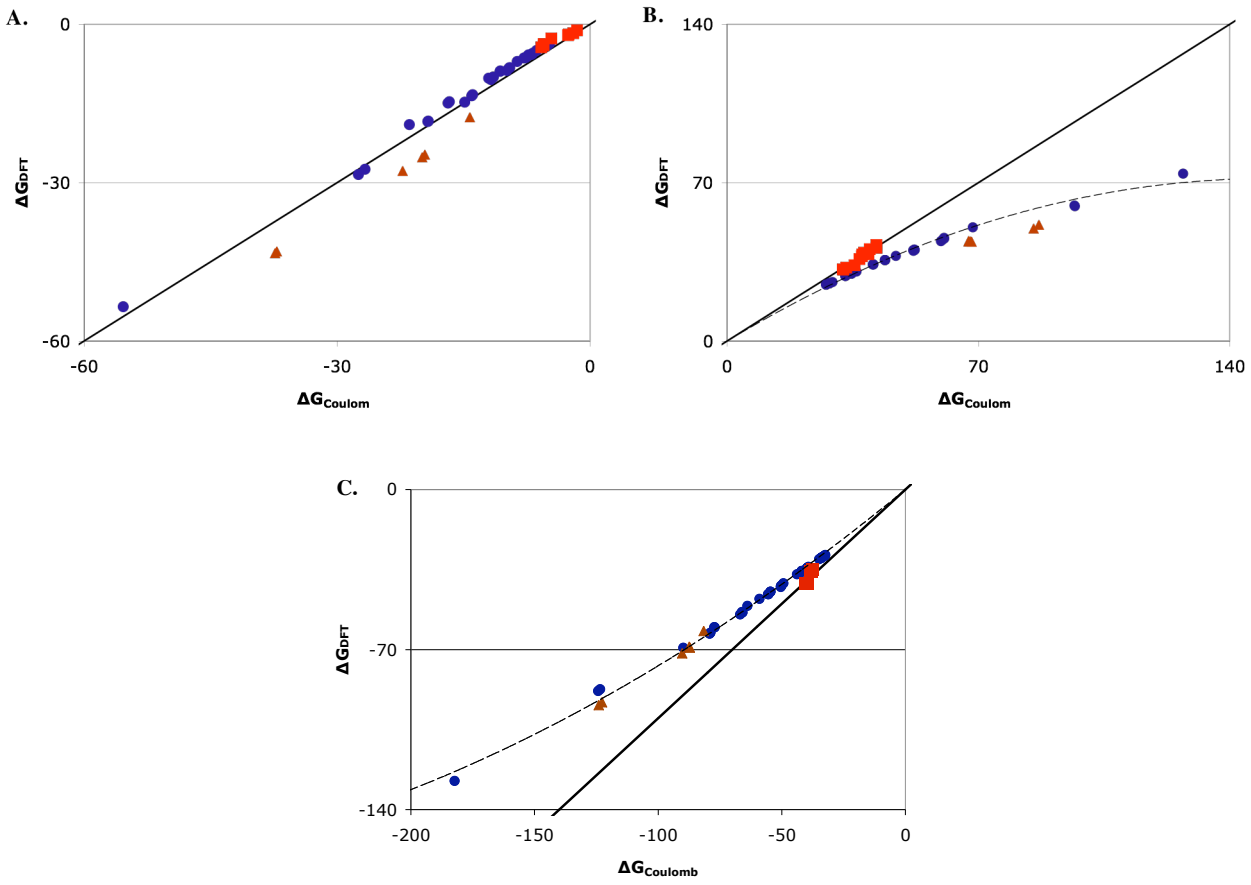

Figure 2.2 Comparison of heme interactions

Comparison of heme interactions with a positive charge (●) near the heme face; (■) in the plane of the heme and near the heme edge; (▲) within 4Å of the heme plane as shown in Fig 2.1 calculated with DFT or CE Coulomb's law with ε=1 for: (A) Neutral, ferric hydroxyl-heme. Ideal straight line through the origin with a slope of 1 is shown. (B) Cationic, ferric water-heme. Line through points is $y = 0.95x - 3.2 \times 10^{-3} x^2$ *is the best fit to a binomial function for the points near the heme face (blue). (C) (Hydroxyl-heme) – (water-heme). Line through pointes is* $y = 0.89x - 1.1 \times 10^{-3} x^2$.

.

3. Computational studies of proton transfer in bacteriorhodopsin using structures from different states.

The pH dependence of the ionization of all acidic and basic residues (Fig 3.1) was determined by MCCE in nine bacteriorhodopsin protein structures trapped in the bR, early and late M states. Nine ionizable residues are buried inside the protein: the Schiff base, Asp 85, 96, 115 and 212, Glu 194 and 204, and Arg 82 and 134. The Schiff base, Asp 85 and 212 are in the center of the protein forming the central cluster. Glu 194 and 204 form the proton release cluster on the extracellular surface. Residue ionization states within these two groups are strongly interdependent, shifting through the reaction cycle as will be described in detail below (Table 1.1, 3.1). All other buried residues maintain the same ionization state in all intermediates. The buried Asp 96 and 115 are neutral from pH 0 to 10. Arg 82 is located between the central and proton release clusters and Arg 134 is near the extracellular surface. Both are deeply buried losing more than 4.5 ΔpK_a units (6.2 kcal/mol) of reaction field energy, which would lower their pK_as to 8.0 in the absence of other interactions with the protein (134). However both remain fully ionized to pH 14. The negatively charged central and proton release clusters stabilize ionization of Arg 82 by more than 12 ΔpK_a units (16.3 kcal/mole). Hydrogen bonds with the backbone carbonyls of Glu 194, Ala 126 and Thr 128 favor ionization of Arg 134 by >9 ΔpK_a units. There are 19 exposed Arg, Lys, Asp and Glu on the cytoplasmatic surface and 4 on the extracellular side. With the exception of Lys30 in 1M0M-1, Glu 74 in 1KG8 and 1C8S and Lys129 in 1DZE, they are all >90% ionized at neutral pH. All Tyr in the protein are neutral at pH 7. The buried Tyr 57, 83 and 185 remain neutral to high pH in all structures.

Central cluster titration. The calculations to be discussed first have all waters deleted and cavities filled by DelPhi with ε=80. Calculations with explicit buried waters will be considered below. The retinal Schiff base (SB) and two acids, Asp 85 and 212 are in the center of bacteriorhodopsin. In all

.

3 bR structures (1C8R, 1C3W and 1KG9) the central cluster has a net charge of –1 (pH 5 - 10) (Fig 3.2) and the Schiff base is 83-97% ionized (pH 0 - 11) (Table 3.1). When the cluster charge is –1 (pH 5 - 10) any residual neutral SB is matched by protonated Asp 85 or 212. The central cluster in 1C8R and 1C3W begins to loose a proton from SB^+ at pH>11, but not in 1KG9. The pK_a for the cluster moving from a charge of 0 to –1 is ≈3 (Table 3.1). 1M0M-1 and 1M0M-2 are resolved in crystals with 40% bR and 60% early M (25). While 1M0M-1 is identified with bR and 1M0M-2 with M, the results of each lie between bR and M. Thus, less SB is ionized in 1M0M-1 than in any pure bR structure while there is more SB^+ in 1M0M-2 than in any M structure (Table 3.1). The results from these structures will be treated separately.

In early M (1DZE and 1KG8), the cluster still has a charge of –1 but the SB is now 88-95% neutral (pH 6 - 11) (Fig 3.3). The cluster proton is shared by Asp 85 and 212 (Table 3.1). The neutral cluster stability has increased, moving its pK_a from 3 in bR to ≈4.5 in early M. At high pH (>11) the cluster begins to lose a proton. Thus, moving from bR to early M, the Schiff base changes from being almost fully ionized to mostly neutral. Asp 85 and 212 also change from being fully ionized to sharing a proton. There is 9% protonated 212 in 1KG8 and 57% in 1DZE. The presence of protonated 212 is unexpected (135) and will be discussed in detail below.

In late M, the SB is ≈90% neutral from pH 7 to 14 (Fig 3.4) with the cluster proton on Asp 85, in agreement with experiment (135). The pK_a of the neutral cluster has shifted up to 5.5 in 1C8S while it remains near 3 in the E204Q mutant 1F4Z (Table 3.1).

Proton release cluster titration. Glu 194 and 204, located near the bacteriorhodopsin periplasmic surface, are implicated in binding the proton to be released to the extracellular space in late M (22,136,137). While, a hydrated hydronium may be the bR proton reservoir (110,138), no hydronium was included in the model tested here. Since adding another binding site would increase the overall proton

.

affinity, these calculations should provide lower limits for the cluster pK_as. The Glu 194 and 204 ionization states are strongly coupled (Fig 3.2-3.4). In all bR and early M structures the cluster has a single proton and a charge of –1 from pH 0 to 10. Glu 204 prefers to be ionized and 194 neutral. In bR there is little proton loss even at pH 14, while in early M ionized 194 begins to be seen by pH 10. The pK_a of 204 remains below pH 0. In the late M (1C8S) 50% of the clusters have a –2 charge by pH 8. Here, the titration is bimodal with some proton loss by pH 5. The late M E204Q mutant (1F4Z) has an Asn at 204. This is replaced by an Asp in the calculation. However, while the central cluster ionization is appropriate for late M, the proton release cluster still binds a proton to pH 10. In addition, the protonated 204 is more stable than in any other structure. Thus the neutral 204, imposed by the mutation, is stabilized by changes that are captured in the crystal structure.

The reduced model of cluster ionization.

The titrations calculated for bR, early and late M structures show shifts in cluster proton distributions (Fig 3.2-3.4). The aim is to understand how the subtle differences between the structures yield these changes. A reduced model was explored which extracts information from the full MCCE treatment. Each of the three central cluster residues have an ionized and neutral conformer yielding 8 microstates: one with a –2 charge, three at –1, three neutral, and one at +1. There are 4 release cluster microstates: two at –1 and one at 0 and one at –2. There are modest differences between cluster ionization in the reduced model and full Monte Carlo analysis especially in late M when residues outside of the cluster change their orientation when the cluster changes ionization state. These errors represents a breakdown of the reduced model mean-field energy approximation,

Why the fully ionized central cluster is favored in the bR state. The three groups in the central cluster are deeply buried. The 2 acids each lose more than 8.5 ΔpK_a units of reaction field (solvation) energy. The SB, with its delocalized charge (and larger Born radius) looses 3.4 ΔpK_a units, which

.

would lower its pK_a from 7, the solution value (111) to 3.6 without other interactions. The rest of the protein favors ionization of each acid by 5-13 ΔpK_a units while destabilizing the ionized Schiff base by 3-5 ΔpK_a units (Eqn. 3). The general positive potential from the backbone (139) contributes 1 to 4 ΔpK_a units to each cluster residue. The buried Arg 82 and 134, and nearby polar groups also raise the potential. With no intra-cluster interactions the acids would be ionized and the SB neutral. In bR, since the acids favor ionization of the SB, and each acid is stabilized by the SB^+ enough to overcome their mutual repulsion, the microstate $SB^+85^-212^-$ is lowest in energy (Fig 3.5A). The other two states with a charge of –1, with the SB neutral and either Asp 85 or 212 ionized, are closest in energy. All microstates with a different net charge are much higher in energy. Thus, if both acids are not ionized the SB will not be charged.

Why the proton shifts from the SB to the cluster acids in the M states. In early and late M the fractionally ionized SB is reduced to 0 - 13% (Table 3.1) primarily because intra-cluster interactions stabilize full ionization less. Retinal isomerization lengthens the distance between the basic SB proton and the acids, reducing the favorable interactions found in bR by ≈0.8 ΔpK_a units. Also, the C15-NZ-CE dipole (δ) is reoriented. When the SB nitrogen points towards the acids as it does in bR, the neutral SB lone pair is destabilized by either ionized acid. When the nitrogen rotates towards the cytoplasm, either ionized acid stabilizes the neutral SB by as much as 3 ΔpK_a units. The non-cluster portions of the protein do stabilize acid ionization less and SB ionization more in M than in bR (changes of ≈1-1.5 ΔpK_a units). In particular, Arg 82 moves towards the extracellular release cluster destabilizing the ionized acids while favoring SB^+. The SB and acids also move into a region with smaller positive backbone potential. Thus, changes in the non-cluster portion of the protein shift the balance between protonated SB and acid little since they stabilizes SB^+ more and either ionized acid less. However intra-cluster interactions destabilize SB^+ and stabilize SB^0, shifting the proton off the Schiff base.

.

What determines the distribution of protonated Asp 85 and 212. While experiments have shown no indication of a protonated 212 in M (140,141), the calculations show $SB^0 85^- 212^0$ and $SB^0 85^0 212^-$ are close in energy especially in early M. Four factors control the balance of ionized 85 and 212 in different structures: Thr 89, Tyr 57 and 185 and the dipole orientation (δ) of the neutral SB C15-NZ-CE. Smaller $|\delta|$ (absolute value of δ) and the Thr 89 hydroxyl favors 85^-, while the two Tyr stabilize 212^- (Table 3.1). The early M structure 1DZE has the most protonated 212. It has a short distance between 85 and The 89 as well as a small $|\delta|$. In late M the proton is bound to Asp 85 in all structures (Fig 3.4, Table 3.1) because $|\delta|$ is larger so SB^0 stabilizes 212^- more (Table 3.1). In late M the shifted F helix moving Tyr 185 and the reoriented Thr 89 also stabilize 212^-.

Why the proton leaves the release cluster in late M. Although Glu 194 and 204 are closer to the surface than Asp 85 and 212, they also loose >8 ΔpK_a units reaction field energy. As in the central cluster, the rest of the protein stabilizes negative charges on the release cluster. The backbone dipoles stabilize each ionized Glu by more than 2 ΔpK_a units. Arg 82 and 134 also stabilize negative charges, as do the nearby Tyr 83 and Ser 193. The result is that in the absence of intra-cluster interactions both acids would have very low pK_as and the microstate $194^- 204^-$ would be fully occupied. It is the mutual repulsion of the charges on the two acids that destabilizes the fully ionized microstate in each intermediate.

In bR, $194^- 204^-$ is disallowed by inter-acid repulsions of 10 to 15 ΔpK_a units in different structures. In addition, the cluster charge:dipole interactions favor the singly ionized state by 2 to 3 ΔpK_a units. The backbone amide interactions with 204^- are ≈2.5 ΔpK_a units more favorable than with 194^-, helping to decide between the two singly ionized microstates. Since the fully neutral state looses interactions with the positive protein potential, it can be above the fully ionized microstate in energy.

.

In early M, Glu 204 has twisted slightly around its CG-CD bond and Glu 194 moves away from the extracellular surface; Arg 82 has moved down towards the release cluster stabilizing ionization of 194 and to a lesser extent 204 and Tyr 83 moves along with the adjacent Arg 82 providing Glu 194 with a stronger hydrogen bond (Fig 3.6B). However, the inter-acid repulsion is still 6-8 ΔpK_a units, while there is still 2-3 ΔpK_a units favorable charge:dipole interaction stabilizing the half ionized cluster. The fully ionized state is therefore still not accessible at pH 7, but is now sufficiently low in energy that it is seen at high pH.

In late M, Glu 194 now points directly up towards the cavity near Arg 82 (Fig 3.6B), while the OE1-CD-OE2 plane of Glu 204 is perpendicular to the membrane normal. At pH 7, the fully ionized state is only 1.3 ΔpK_a units higher in energy than 194^0204^- so now the fully ionized release cluster is seen at neutral pH. The most important reason is that the 2 acids have moved apart. This reduces the inter-acid repulsion to be $\approx$6 ΔpK_a units and the favorable charge:dipole interactions to be less than 1ΔpK_a unit. There are also small changes in interaction with Arg 82 and the backbone. In the E204Q mutant 1F4Z structure, the release cluster remains protonated. Arg has moved down, but 194 has not moved up so it is still too close to 204 for both to be ionized. As the protein is trapped with a neutral 204 the preferred state has 194^- rather than 204^- because Arg 82 is positioned to stabilize 194^-, while in the other late M structure, with a wild type release cluster, both acids have similar interactions with the base.

Coupling between the central and release clusters.

The M state is characterized by a shift of the proton from the SB^+ to Asp 85^-, maintaining a central cluster charge of –1. There is no direct electrostatic coupling between this change in charge distribution and release cluster deprotonation. Thus, the early M structures and E204Q late M structure have a protonated release cluster and a neutral SB. Arg 82 does facilitate coupling between the

.

clusters. The $SB^0 85^0 212^-$ central cluster has a less favorable interaction with the Arg than does $SB^+ 85^- 212^-$. This allows the Arg to move towards the release cluster stabilizing a –2 release cluster charge. However Arg movement does not ensure proton release as they are in early M and E204Q structures when the two acids in the proton release cluster are still too close. In late M 194 and 204 move apart lowering the pK_a for proton release.

While the two clusters are separated by 12 Å, (between 85:212 and 194:204 carboxylates) the interaction between their net charges is 2 ΔpK_a units (2.7 kcal/mol) providing direct coupling between their total ionization. Some of the differences between the titrations calculated with different structures trapped in the same state reflect this. For example, in bR structures 1C8R and 1C3W the central cluster begins to loose a proton, moving to a charge of –2, at pH >10 while the release cluster retains its –1 charge to pH >14. In 1KG9 the results are reversed as the exit cluster begins to be deprotonated and the central cluster keeps its –1 charge. The total charge on the two clusters is quite similar in all structures (Fig 3.2). Thus, these pK_as in the 2 clusters are close. Small structural differences change the ordering of the cluster pK_as. Their interaction then ensures that ionization of one cluster suppresses the proton release from the other.

The 1C8S (late M), titration highlights the coupling of the cluster net charges (Fig 3.4). The central cluster 0 to –1 and proton release cluster –1 to –2 titrations are now bimodal with pK_as at 4.5 and 7.5 as the two clusters move from a net charge of –1 (pH < 4) to –3 (pH >8). The break in the global titration at pH 6 occurs when the total charge is –2, with ≈74% of the central cluster and 26% of the proton release cluster ionized. The cluster pK_as in M differ from those in bR because Arg 82 motion now stabilizes the central cluster 0 charge state. The release cluster –2 state is also stabilized by the Arg motion as well as by separation of its two acids.

Addition of explicit waters to cavities inside the protein.

.

Calculations with explicit waters. MCCE uses a low, continuum dielectric constant for the protein of 4, but allows polar group reorientation which stabilizes the equilibrium charge distribution. This provides an effective dielectric greater than 4 and yields information about the atomic details of the response. When there are cavities in a protein, DelPhi fills them with high dielectric (ε=80). Nearby charges are now stabilized by the continuum reaction field energy which models an averaged water orientation. Calculations were initially carried out without explicit waters because MCCE calculations are sensitive to the pre-selection of starting conformers. If a conformer capable of a favorable hydrogen bond is not included, charge state energies may be wrong. This is a particular problem with waters which have translational and rotational degrees of freedom which need to be carefully considered when assigning initial positions.

MCCE titrations were calculated in bR (1C8R), early (1DZE) and late M (1C8S) intermediates with explicit waters in cavities. There are 13 internal water oxygen positions in the bR, 10 in the early M and 10 in late M structure. Additional water oxygens were added as described in Methods. Each oxygen has protons placed to optimize local hydrogen bonding opportunities, resulting in 896, 279, and 313 water conformers for the 66, 36, and 34 oxygen positions tested in bR, early and late M. In Monte Carlo sampling waters have a conformer that has left the protein. This is chosen if the interaction of an individual water with bulk solvent (–1.3 ΔpK_a unit) is more favorable than those found in this binding site (142). Cavities were overfilled, including oxygens separated by small translations which will compete during Monte Carlo sampling. There are on average 17.5 water molecules bound in bR, 12.8 in early and 7.8 in late M.

Titration with explicit water. For both clusters the calculated titrations are similar with continuum or explicit waters in cavities (Fig 3.2-3.4). The total central cluster charge is independent of the water model at low and intermediate pH. It retains its single proton somewhat better at high pH with

.

explicit waters. There are some changes in the residue ionization states. At low pH in bR the proton is exclusively bound by Asp 85, in better agreement with the experimental estimates of the pK_as of the 2 Asps (140,143,144). Above pH 9.5 Asp 85 ionization increases from 92% to 98%, as the cluster as a whole begins to deprotonate. This may capture the small secondary wave in the Asp 85 titrations near pH 9.5 found experimentally (21,145). This deprotonation is coupled to proton loss from Asp 115. At the plateau of this wave of titration the central cluster has an average charge of –1.07 and Asp 115 of –0.93. In early M explicit waters modify the distribution of protonated 212 and 85 slightly at all pHs. In late M at higher pH explicit waters stabilize $85^0212^-SB^0$ so no occupied microstates have ionized SB or Asp 85.

There are modest changes in release cluster titrations when the cavities are filled with explicit rather than continuum water (Fig 3.2-3.4). There are no changes in the bR state. In early M the cluster is more fully deprotonated at high pH. In late M the fully deprotonated cluster is more stable at lower pH as explicit water provides additional stabilization of the 194^-204^- state. This lower pK_a is closer to the experimentally determined value of 5.8 (137). Long-range coupling, discussed above, yields the bimodal titration curves.

Reduced model with explicit water. When explicit waters fill cavities residues loose favorable continuum reaction field energy but gain favorable pair-wise interactions with water. The waters rearrange to stabilize the low energy ionization states. In general, the most occupied microstate found in the continuum cavity analysis is stabilized most, as seen in the reduction of minor ionization states in the titrations (Fig 3.2-3.3).

Waters in the cavities defined by 402 and 401/406 in the bR structure 1C8R are the only ones that influence the central cluster (Fig 3.6). In bR waters support ionization of all 3 cluster groups, although long-range interactions with the buried Arg 82 and 134 destabilize SB^+ more when interior cavities are filled with explicit waters with ε=4 rather than continuum water with ε=80. The water

.

cluster is closer to Asp 85 than 212, and so stabilizes 85^- by about 2 ΔpK_a units and Asp 212^- by ≈1 ΔpK_a units. An added water near 402 is a hydrogen bond donor to both 85 and 212, stabilizing their ionization by 1.5~1.7 ΔpK_a units. This lowers the energy of microstates with ionized 85 and 212. In early M, two waters stabilize Asp 85 ionization by 2.7 ΔpK_a units without influencing Asp 212^-. Asp 85 is now ≈5% more ionized and Asp 212 10% more protonated than with continuum water. In late M a chain from the protonated Asp 85, through two waters in the 401/406 cluster to 212^- stabilize the microstate with Asp 85 neutral and 212 ionized by 3.4 ΔpK_a units.

The switch from the averaged water response encapsulated in ε=80 to explicit, movable waters changes the release cluster behavior in modest ways. In bR, a water added near 403 stabilizes both Glu 194^- and 204^- by 1.5 to 2 ΔpK_a units but this is still insufficient to overcome the inter-acid repulsion. In early M, water stabilizes ionization of both residues by a small amount, increasing cluster ionization at high pH. The water changes microstate energies less in early M than in bR, but since the state energies are closer together it is easier to modify the site titrations. In late M, two waters stabilize ionization of 194 and 204 by ≈2 ΔpK_a units. Two others are oriented to stabilize194^-204^- at high pH by 3-4 pK_a units while they reorient to favor the 194^-204^0 microstate at low pH.

.

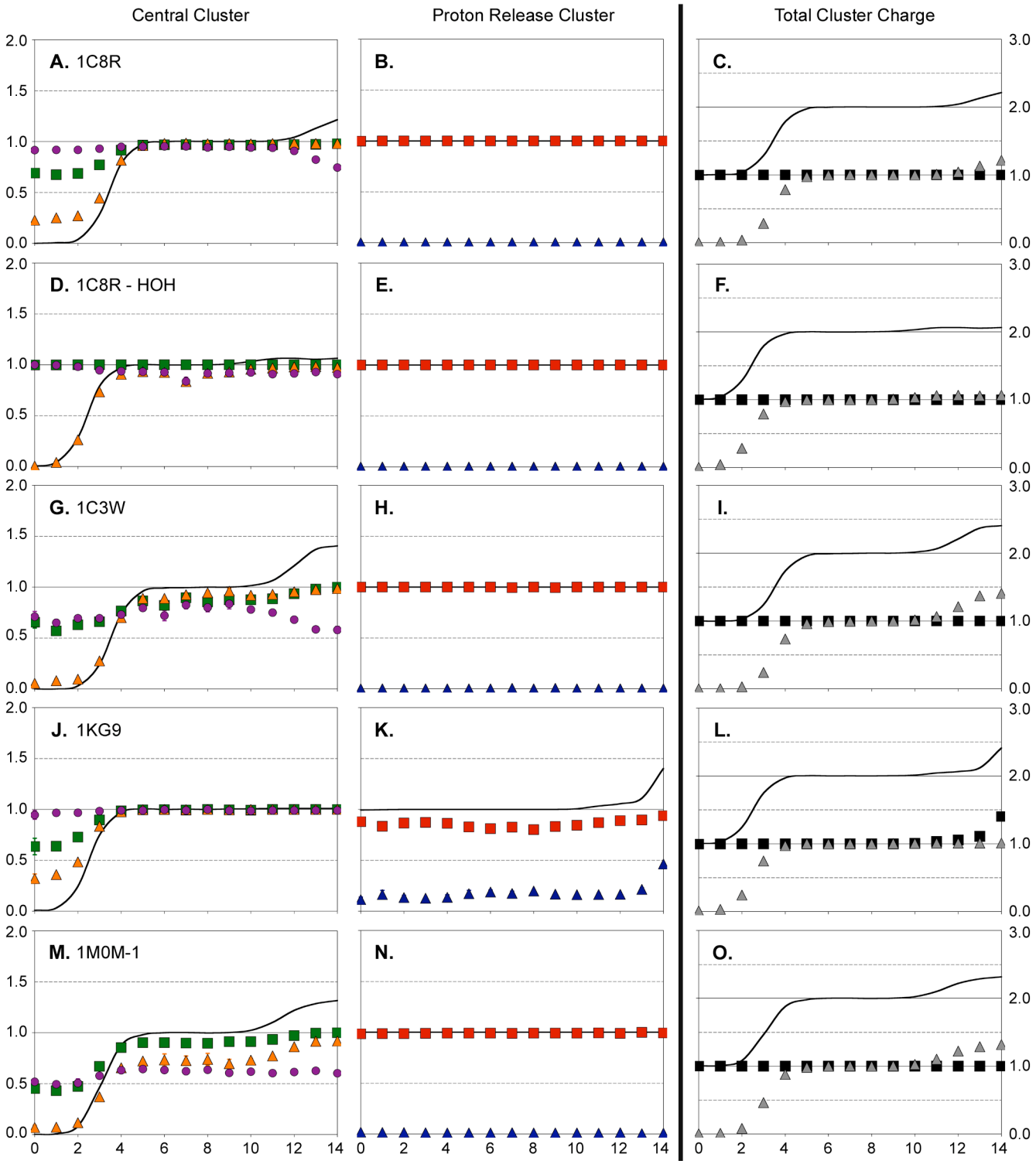
Central Cluster
Proton Release Cluster
Total Cluster Charge
A. 1C8R
B.
C.
D. 1C8R - HOH
E.
F.
G. 1C3W
H.
I.
J. 1KG9
K.
L.
M. 1M0M-1
N.
O.

Figure 3.1 Calculated bR titrations

MCCE calculated bR ionization states of (A, D, G, J, M) central and (B, E, H, K, N) release cluster. (C, F, I, L, O) sum of charges on the two clusters in (A-F) 1C8R, (G-I) 1C3W, (J-L) 1KG9 and (M-O) 1M0M-1. Calculation with protein cavities filled by (A-C, G-O) continuum and (D-F) explicit water. (A, D, G, J, M) Central cluster: Line=–1 x net charge; ionization (●) SB, (▲) Asp85 and (■) 212. (B, E, H, K, N) Release cluster: Line = –1 x net charge. Ionization (▲) Glu 194 and (■) 204. (C, F, I, L, O) Line = –1 x sum of total charges on the two clusters; -1 x cluster charge in (▲) central and (■) release cluster. Values and their error bars represent average and standard deviation of five rounds of Monte Carlo sampling.

.

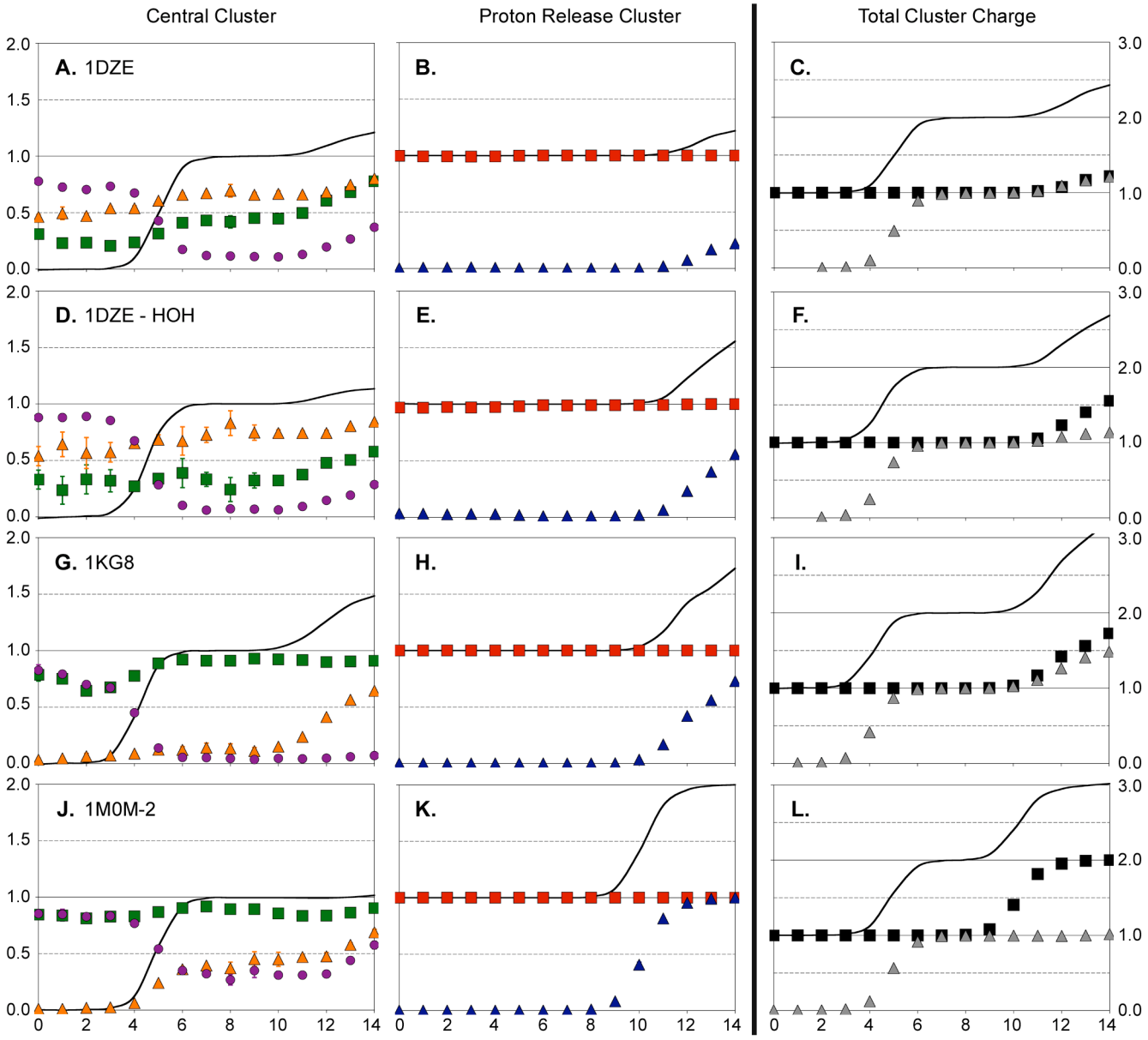

Figure 3.2 Calculated early M titrations

MCCE calculated ionization states in early M of (A, D, G, J) central and (B, E, H, K) release cluster. (C, F, I, L) sum of total charges on the two clusters in (A-F) 1DZE, (G-I) 1KG8 and (J-L) 1M0M-2. Calculation with protein cavities filled by (A-C, G-L) continuum and (D-F) explicit water. Lines and symbols that same as in Fig 3.1.

.

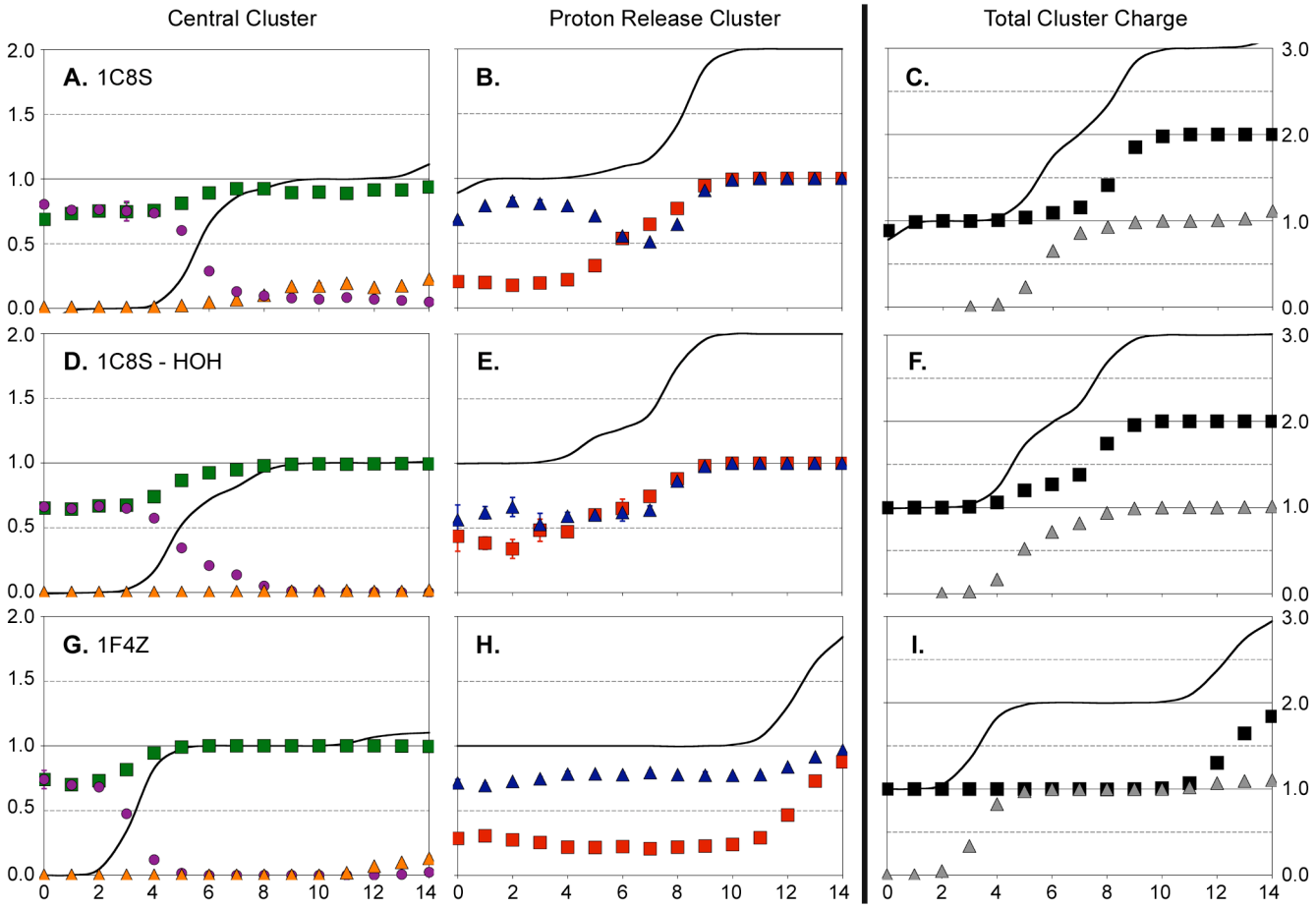

Figure 3.3 Calculated late M titrations

MCCE calculated late M ionization states of (A, D, G) central and (B, E, H) proton release cluster and (C, F, I) sum of total charges on the two clusters in (A-F) 1C8S and (G-I) 1F4Z. Calculation with protein cavities filled by (A-C, G-I) continuum and (D-F) explicit water. Lines and symbols that same as in Fig 3.1.

.

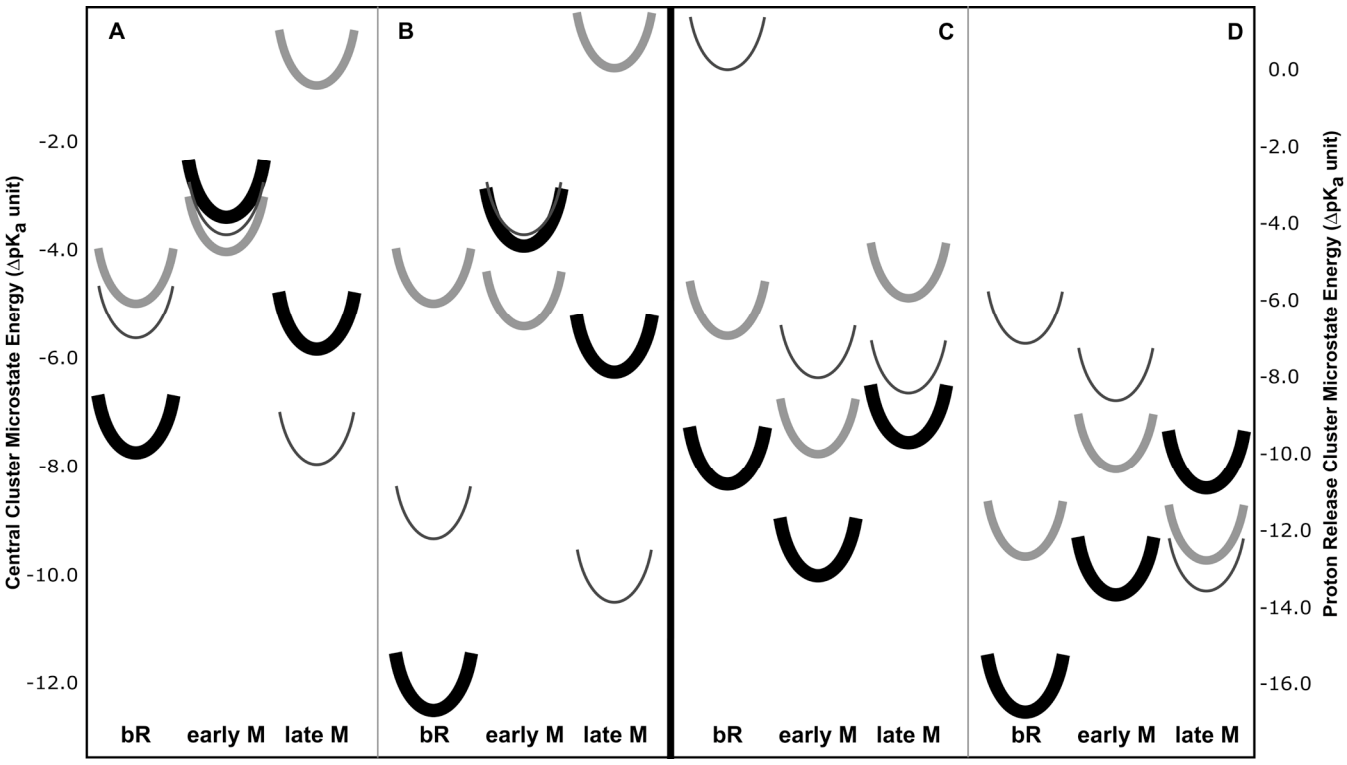

Figure 3.4 Reduced model energy levels of the three central cluster microstates with a charge of –1 and the low energy proton release cluster microstates

(A, B) Central cluster microstate energies of (thick) $SB^{+}85^{-}212^{-}$, (medium) $SB^{0}85^{-}212^{0}$, and (thin) $SB^{0}85^{0}212^{-}$. (C, D) Proton release cluster (thick) $E194^{0}E204^{-}$, (medium) $E194^{-}E204^{0}$ and (thin) $E194^{-}E204^{-}$. pH is 7. (A, C) continuum and (B, D) explicit water calculations. Each microstate is represented as an arbitrary shaped harmonic energy well. The minimum energy is derived from the reduced model.

.

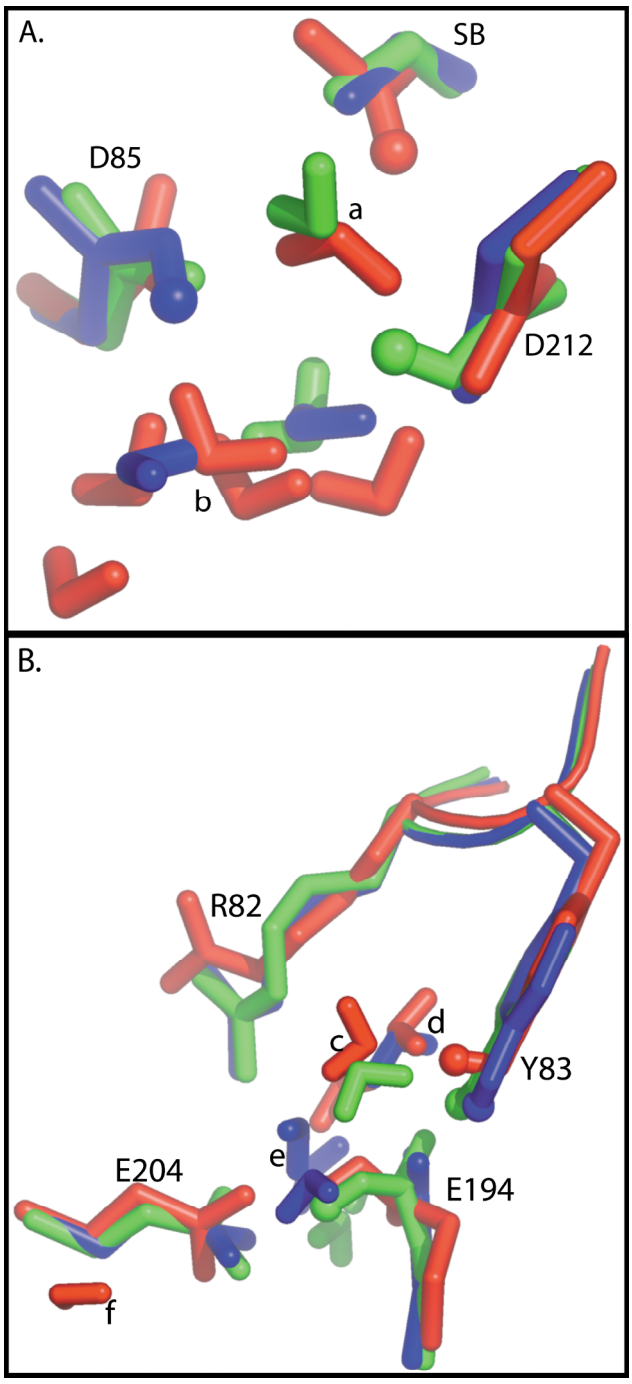

Figure 3.5 Water molecules in central and proton release cluster

In (A) central and (B) proton release clusters, waters reorient to stabilize ionization states in (red) bR, (green) early and (blue) late M states. Waters with occupancy of >40% in Monte Carlo sampling are

.

shown. Thinner line in B: backbone connecting R82 and Y83. The active, transferred proton on cluster residues and the Y83 hydroxyl proton are small spheres. The crystallographic waters often have unfavorable Lennard-Jones interactions with the protein and so are not retained in Monte Carlo sampling. Waters were added to provide alternate positions and to fill cavities. The numbers refer to the nearest crystallographic waters.

(A) Central cluster: (a) Water 402 between SB, Asp85 and 212 is 97% occupied in bR and early M stabilizing the proton on SB in bR and on D212 in early M. The cavity is not big enough for water in late M. (b) Water cluster 401/406 (bR, late M), 603 (early M) lies to the extracellular side of D85 and D212. Waters were added to fill cavity. The average cavity water occupancy is 1.7-1.9 in bR, 0.61 in early, and 1.9 in late M. The cavity is smaller in early M. This cluster can rearrange to stabilize ionization of either acid. In late M the waters form a hydrogen bond chain from D85 through W401/406 to $D212^{-}$.

(B) Proton release cluster (c) Added waters between E194 and Y83 in bR and early M. One accepts a hydrogen bond from Y83 in bR. In early M, Y83 moves downward along with Arg 82 to hydrogen bond to E194 reducing the Y83-water interaction. (d) Water cluster 403/405 and 404 (bR only) with additional waters added to fill cavity lies between R82, E194 and E204 in bR and late M. Occupied conformers accept a hydrogen bond from Arg 82 and donate one to Glu 204. (e) Waters 608 (early M) and 404 (late M) and additional waters lie between Glu 194 and 204. The average water occupancy in the cavity is 1 in early and 3 in late M. This cavity is closed in bR. One water is a hydrogen bond acceptor from Arg 82 and donor to Glu 204, another is a hydrogen bond donor to both Glu 194 and 204. (f) Water 944 is added to the extra-cellular side of the cluster in bR. It makes a hydrogen bond to E204.

.

	PDB	Resolution (Å)	PDB author	Fraction Ionization pH 7					pH dependence of summed charge of both clusters		δ
				SB	Asp85	Asp212	Glu194	Glu204	Lower pK_a	Higher pK_a	
bR	1C8R	2.3	Luecke et al.	0.95	0.98	0.97	0.00	1.00	3.5	13.5*	-61.6
	1C3W	1.55	Luecke et al.	0.83	0.93	0.90	0.01	1.00	3.5	12.0*	-69.8
	1KG9	1.81	Facciotti et al.	0.99	1.00	0.99	0.17	0.83	2.5	11.1*	-65.0
	1M0M-1	1.43	Lanyi and Schobert	0.62	0.73	0.90	0.01	0.99	3.1	11.5*	-64.3
early M	1DZE	2.5	Takeda et al.	0.12	0.67	0.43	0.00	1.00	5.0	12.6*	75.9
	1KG8	2.0	Facciotti et al.	0.05	0.14	0.91	0.00	1.00	4.2	11.6	99.5
	1M0M-2	1.43	Lanyi and Schobert	0.32	0.39	0.92	0.00	1.00	4.8	10.2	-51.3
late M	1C8S	2.0	Luecke et al.	0.13	0.07	0.92	0.51	0.65	5.5	8.3	102.8
	1F4Z	1.8	Luecke et al.	0.00	0.00	1.00	0.79	0.21	3.3	12.4	122.6

Table 3.1 Fraction cluster residue ionization at pH 7, pK_as for total cluster ionization and Schiff base orientation in each structure

Lower and higher pK_a from data in Fig 3.1-3.3. *: pK_a of total ionization of two clusters plus Asp115. δ: angle defining orientation of SB relative to Asp85 and 212.

.

4. Modeling proton transfer in bacteriorhodopsin using one single structure.

The aim of this work is to model different intermediates of the bacteriorhodopsin reaction cycle using a single ground state structure (1C8R (30)). Conformational changes that are proposed to couple to the reaction are modeled by additional conformers in MCCE. Different intermediates are defined by constraints on conformer selection during the Monte Carlo sampling. Calculations of ground state bacteriorhodopsin are carried out with retinal fixed in all-*trans* configuration. A state that mimics the L intermediate is modeled by fixing retinal in 13-*cis* configuration. The next intermediate is calculated by keeping retinal fixed at 13-cis when moving Arg 82 closer to the proton release cluster. This mimics the early M intermediate. And finally an additional constraint is applied to Glu 194 to mimic the late M intermediate. In each case all other residues are allowed to sample all their conformers freely.

The earlier bacteriorhodopsin calculations used MCCE 1.0 (Chapter 3). This new study was carried out with the additional conformational flexibility of MCCE2 (104). The total charge on each cluster at the neutral pH is consistent with the earlier studies, however the charge distribution on the proton release cluster, and the pH-dependence of the cluster changes (Fig 4.1).

In the central cluster, the Schiff base, Asp 85 and 212 are all fully ionized at neutral pH (Fig 4.1). Asp 85 has a pK_a of 3, while the Schiff base and Asp 212 both stay fully ionized between pH 0 and 14. These results are essentially the same as the earlier studies (Chapter 3) and in good agreement of the experiments (140,143,144,146).

Negative charge of the proton release cluster is stabilized by the interactions with backbone, Arg 82 and 134, Tyr 83 and Ser 193. However the large intra-cluster interaction between the two glutamic acids disfavors both being ionized. This leads to -1 total charge on the cluster, in agreement with earlier calculations (114). However the negative charge on the cluster is not localized on one group as found in

.

the earlier studies (114). Instead it is distributed to give both Glu 194 and 204 50% ionization. This shift is due to the additional structural flexibility given to Glu 194 in the new MCCE. In the crystal configuration, ionized Glu 204^- is favored over Glu 194^- by the backbone and nearby polar residues. In this work, Glu 194^- can change conformation to receive better interactions with backbone, Ser 193 and Glu 204^0, allowing Glu 194^- and Glu 204^- to be about equally stabilized. A more distributed proton on this cluster is in better agreement with the suggestion that this proton is stored in a hydrogen-bonded network (24,138,144,147-150). At high pH, the proton on the proton release cluster is released with a pK_a of 10.5 in the ground state structure. Compared to the earlier calculations where this proton is not release at pH>14, this result is in better agreement with the measured pK_a value of 9.5 (145,151). If both glutamic acids are fixed in the crystal conformation, the large unfavorable interaction of 13ΔpK units keeps the proton bound very tightly. The additional MCCE conformations of Glu 194 reduces this interaction to 8ΔpK units, lowering the energy of the doubly ionized cluster, allowing it to be accessible when pH increases.

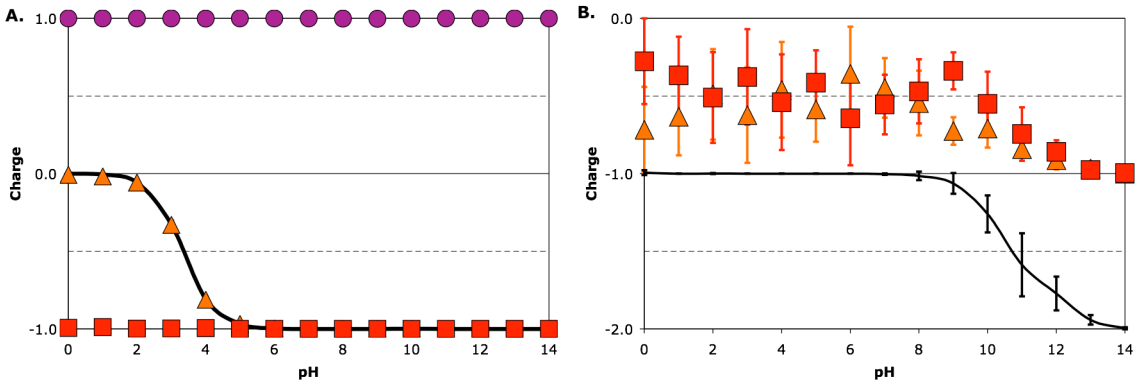

Figure 4.1 MCCE calculated ionization states in the ground state

(A) Central cluster: Line: total cluster charge; fraction of charge on (●) SB, (▲) Asp85 and (■) 212. (B) Proton release cluster: Line: total cluster charge; fraction of charge on (▲) Glu 194 and (■) 204.

.

Values and their error bars represent average and standard deviation of ten rounds of Monte Carlo sampling.

L intermediate model

Retinal isomerization is the first structural change after light is absorbed. The modeled L structure uses an isomerized retinal and the other amino acids can sample all conformations freely. The essential structural changes are localized in the retinal binding pocket (26,28). Compared to the ground state, the changes of ionization states are in the central cluster, while no change is found in the proton release cluster (Fig 4.2). Due to the isomerization, the charge on SB^+ moves away from Asp 85 and 212, destabilizing $SB^+85^-212^-$ state. Meanwhile, the neutral SB dipole flips around, stabilizing $SB^085^0212^-$ and $SB^085^-212^0$ states by changing the unfavorable interaction between the neutral SB dipole and either acid to be favorable. This shifts the balance of the proton on SB^+ and Asp 85^0 or Asp 212^0. Only a small ionization shift of 0.16 proton from the SB to Asp 85 is observed. The changes of interactions driving the proton shift has been identified in the earlier studies (114). In addition, due to the isomerization, the interaction between SB^+ and a proton on Asp 85 or 212 is reduced. Therefore the pK_a for moving the total charge on the cluster from 0 to -1 is shifted higher by about 0.5 pK unit, stabilizing the proton uptake.

The proton shift found in this calculation is smaller than the earlier calculations (114), where over 90% of proton is shifted using early M structures. The difference is mostly from the structure of 13-*cis* retinal. Isomerization of the retinal is modeled here in CHARMM minimization subroutine with the rest of the protein in the ground state. As a result the SB is found to be closer to Asp 85 and 212 than in the M structures. Therefore the changes of the interactions between SB^+ and the two aspartic acids are

.

smaller. Conformational changes of Arg 82 found in the early M structures also contribute to the proton shift in the central cluster. They are modeled is the next calculation.

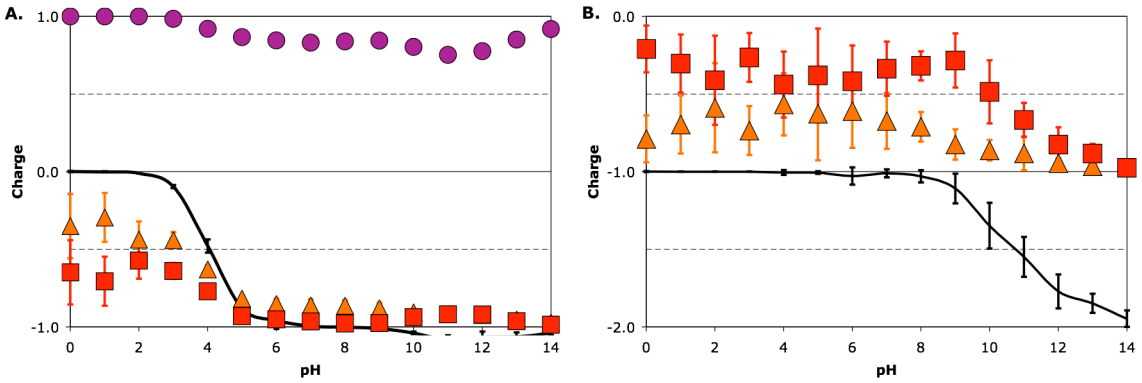

Figure 4.2 MCCE calculated ionization states in the modeled L intermediate

(A) Central cluster: Line: total cluster charge; fraction of charge on (●) SB, (▲) Asp85 and (■) 212. (B) Proton release cluster: Line: total cluster charge; fraction of charge on (▲) Glu 194 and (■) 204.

Early M intermediate model

In the next state, Arg 82 is fixed to face the extracellular side. This state is proposed to model the early M intermediate. In the early M crystal structure (29), conformational changes of Arg 82 have been observed. Changes of ionization states are observed in both central and proton release cluster. In the central cluster, now about 0.5 proton is shifted from the SB to Asp 85 and 212. This proton transfer has been shown by experiments (135) and also found in the earlier calculations (114,152). The additional proton shift is caused by the reduced interactions between the central cluster and Arg 82, which favors Asp 85^{-} and disfavors SB^{+}.

The total charge is also changed on both clusters at pH 7. In the central cluster, there is a fraction of proton uptake (0.15). Meanwhile, a fraction of proton is released from the proton release cluster (0.20). Conformational change of Arg 82 favors the proton release on the release cluster by 3ΔpK units,

.

shifting the pKa down from 10.5 in the ground state. Proton uptake on the central cluster is also favored 1.5 ΔpK units by this conformational change. In addition, the proton release and proton uptake are coupled by 2.5ΔpK units of long-range interaction. This leads to the bimodal titration of the total charge on both clusters.

Although about 50% protein still stays in the ground state, the rest of the protein is in a mixture of later intermediates. About 30% of the protein is now stabilized in early M state, 5% in late M state and 15% in O state. In the earlier studies using early M state crystal structure (114), most of the protein is localized in early M ionization state. The additional conformational flexibility in this study allows doubly ionized proton release cluster to be stabilized so there is more distribution in late M and O ionization states.

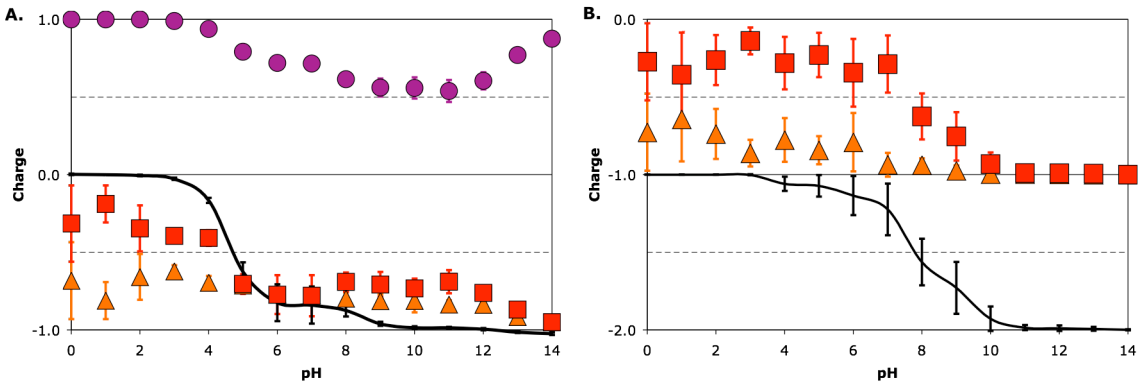

Figure 4.3 MCCE calculated ionization states in the modeled early M intermediate

(A) Central cluster: Line: total cluster charge; fraction of charge on (●) SB, (▲) Asp85 and (■) 212. (B) Proton release cluster: Line: total cluster charge; fraction of charge on (▲) Glu 194 and (■) 204.

Late M intermediate model

The next step is to constrain the conformation of Glu 194. This conformational change is observed in the crystal structure trapped in the late M structure (153). It is the key to stabilize the doubly

.

ionized proton release cluster as observed here at high pH in the ground state and in the earlier studies (114,152).

Large shift of cluster charge pK_a is observed on both clusters. In the central cluster, the proton uptake now has a pK_a of 7.3 and is now 65% stabilized with a net charge of 0 at pH 7, in good agreement with the pK_a of 8.2 measured in the late M state of D96N mutant (154). In the proton release cluster, the proton release is now fully stabilized at neutral pH and the pK_a for formation of the -2 charged state is shifted down to 3.5. The calculated pK_a is lower than the measured value of 5.8 in M state (137). Since this calculation forces Glu 194 to take a conformation stabilizing the proton release, the calculated low pK_a is expected.

The calculated ionization states here give 35% of the protein in the late M and 65% in the O intermediate. Compared to the earlier calculation using late M state structure (114), where 84% of the protein is in early M ionization states, 4% in late M and 12% in O, the later intermediates are more stabilized. This is mostly due to the difference of the conformation chosen by Arg 82 and Glu 194. A strong salt bridge is found here formed by these two residues with 11ΔpK units of favorable interaction. This is 3~4 ΔpK units stronger than in the earlier calculations, stabilizing the proton release.

.

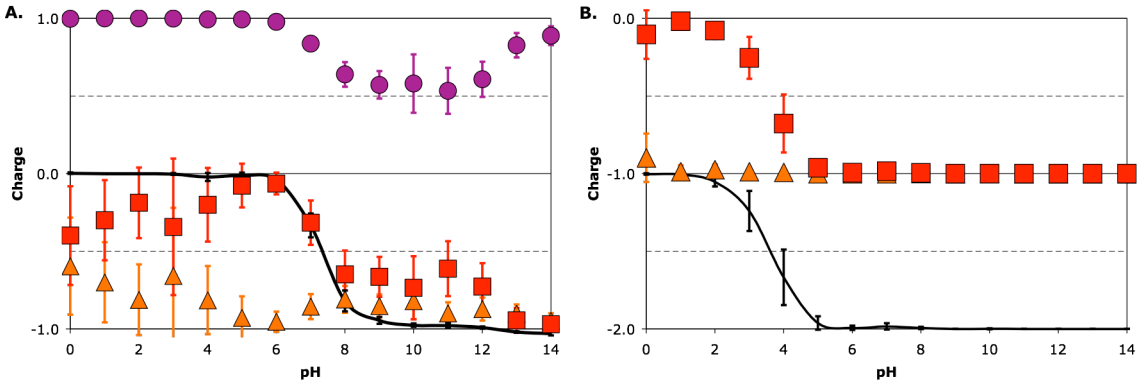

Figure 4.4 MCCE calculated ionization states in the modeled late M intermediate

(A) Central cluster: Line: total cluster charge; fraction of charge on (●) SB, (▲) Asp85 and (■) 212. (B) Proton release cluster: Line: total cluster charge; fraction of charge on (▲) Glu 194 and (■) 204.

Conclusion

In this work, the pH-dependence of ionization states in bacteriorhodopsin is calculated with MCCE. Using one ground state structure and by applying three key structural changes, calculations showed sidechain conformation and charge distribution equilibrate around those key changes, stabilizing ionization shifts in the right direction. In the ground state, one proton is localized on the SB in the central cluster and one proton distributed in the proton release cluster. When isomerization of retinal occurs in L intermediate, there is a little proton transfer from the SB to Asp 85 and 212. Then as Arg 82 moves from the cytoplasmic side to the extracellular side in early M, proton transfer in the central cluster is more stabilized. Proton release from the proton release cluster and proton uptake on the central cluster is partially observed in the early M model and more stabilized in late M as Glu 194 moving away from Glu 204. Most of the proton shift found in the earlier calculations with structures trapped in different intermediates can be reproduced in using one ground state structure. Although often

.

with less proton transfer. In addition, with the additional structural flexibility given by MCCE2, late M and O intermediates are more stabilized with the conformational changes.

.

5. Calculations of cytochrome c oxidase

Ionization states of all residues in the three subunit *Rb. sphaeroides* cytochrome c oxidase were calculated with Multi-Conformation Continuum Electrostatics (MCCE) in seven different anoxygenic redox states (35): fully oxidized (OOOO, O state), singly reduced states with Cu_A reduced (ROOO), Heme a reduced (OROO) or Cu_B reduced (OOOR, E state), mixed valence states with Heme a and Cu_B (ORORO) or both Heme a_3 and Cu_B (OORR, R state) reduced, and the fully reduced state (RRRR). Sites where proton uptake or internal transfer is coupled to electron transfer are identified by the calculated differences in residue protonation in different redox states (Table 5.1). In addition, $pK_7^{'}$ (Eqn. 7) was calculated for key residues, which provides the energy required to move away from the equilibrium protonation state at pH 7 (Table 5.2).

Residues responding to the redox state changes can be assigned to three clusters. The BNC cluster encompasses hydroxyl groups coordinated to Heme a_3 and/or Cu_B (65-67). In addition, the Heme a and a_3 propionic acids (155), Glu I–286 (39-41), Tyr I–288, Lys I-362, and His I–334 (71,72) are essential ionizable residues in the vicinity of the BNC, which are in a position to couple proton uptake to cofactor reduction (Fig 5.2). On the proton release side of the protein, His I–93 and Glu I–182 form a cluster which changes ionization as the oxidase cofactors change charge (Fig 5.3a). The third cluster which includes His I–127 and Glu I–539 is found on the proton entry side near the D-channel Asp I–132 (Fig 5.3b).

.

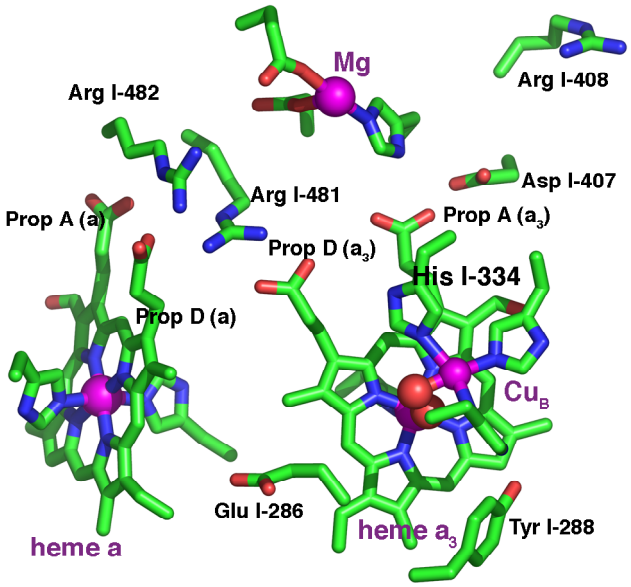

Figure 5.1. Heme a and a_3, Cu_B, the Mg cluster and nearby ionizable residues

Mg and cofactor metal atoms are purple spheres. The oxygens of the BNC water ligands are red. The Heme a_3 water oxygen is behind that of Cu_B. The only labeled cofactor or ion ligand is His I–334, a Cu_B ligand whose pK_a is studied here.

BNC hydroxyl pK_as. In the absence of oxygen, Cu_B and Heme a_3 each binds one water or hydroxide as their 4th and 6th ligand, respectively. To couple proton uptake to BNC reduction, the hydroxyl pK_a must be lower than the solution pH in the fully oxidized state, and higher following reduction. The $pK_{a,sol}$ for deprotonation of an oxidized His-aquo-Heme has been measured to be 9.6 (126), while that of an oxidized Cu(II) with 3 His ligands is estimated to be 9.4 (128,129). In the fully oxidized protein, MCCE Monte Carlo sampling places a hydroxyl on Cu_B down to a pH below 4, and a water on Heme a_3 beyond pH 11. Using the Boltzmann sampled ionization states of all other residues at pH 7 (Table 5.1), the $pK_7^{'}$ of –9.4 for aquo-Cu_B indicates that 16.4 ΔpK units (22.3 kcal/mol) are

.

needed to protonate the hydroxyl at pH 7 (Eqn. 6, Table 5.2). The Heme a_3 $pK_7^{'}$ is 10.7, indicating the cost of removing a proton from this water is 3.7 ΔpK units (5.0 kcal/mol).

The factors that determine $pK_7^{'}$ can be analyzed using Eqn. 4 (Table 5.3). The oxidized hydroxyl-Cu_B has a net charge of +1. When a water is bound its charge is +2. The oxidized hydroxyl-Heme a_3 has a net charge 0 and it is +1 when a water is bound. The larger charge means that the water-complexes loose more solvation energy ($\Delta\Delta G_{rxn}$) than the hydroxyl-complexes when they are moved into the protein. The difference stabilizes the hydroxyl form, lowering the in situ aquo-cofactor $pK_7^{'}$. The $\Delta\Delta G_{rxn}$ lowers the Heme a_3 $pK_7^{'}$ by only 3.9 pH units, while it destabilizes the water-Cu_B by 14.7 ΔpK units (Table 5.3). The favorable interaction of each hydroxyl with the positive charge on the other BNC metal center lowers the $pK_7^{'}$ by >10 ΔpK units. The interaction of the hydroxyl with the rest of the protein (ΔG_{pol} and $\Delta G_{res,7}^{mfe}$) raises its $pK_7^{'}$ by ≈6 ΔpK units (Table 5.3). The $pK_7^{'}$ of aquo–Cu_B is –9.4 when a water is bound to Heme a_3 (Table 5.2), while the aquo–Heme a_3 $pK_7^{'}$ is –5.0 with water on Cu_B. As a result, at least one hydroxyl is very stable in the oxidized BNC. However, the repulsion between the two hydroxyls destabilizes the ionization of both by 15.7 ΔpK units. This is consistent with earlier MCCE1.0 calculations on *P. denitrificans*, which also found one hydroxyl group in the BNC in the fully oxidized enzyme (66). Cu_B with the lower $pK_7^{'}$ binds the hydroxyl, while a water is bound to Heme a_3. This is in agreement with the optical absorbance measurements, which have shown a high-spin Heme a_3 in oxidized cytochrome c oxidase (156). Hydroxyl-hemes are expected to be in a low-spin state (67,130), while water-hemes are generally in a high-spin state (67). EXAFS and ENDOR experiments support the presence of a hydroxyl-Cu_B in the oxidized enzyme (65).

When Cu_B is reduced forming the OOOR state, a single hydroxide is calculated to remain in the BNC. A $pK_{a,sol}$ for an aquo-Cu(I) Cu_B analogue is not found in the literature. The reduced Cu_B, with its

.

smaller charge, should have a higher aquo-heme pK_a then the Cu(II) complex. The calculation shows even leaving the $pK_{a,sol}$ for the reduced hydroxyl-Cu_B at 9.4, the hydroxyl in the BNC shifts to the oxidized Heme a_3. The desolvation penalty, which strongly destabilizes the water-Cu_B(II), now favors water-Cu_B(I) (Table 5.3). Although, the reduced water-Cu_B(I) complex has the larger net charge, the hydroxyl complex has a large dipole moment, and more reaction field energy in water. The $pK_7^{'}$ of hydroxyl-Cu_B is 12.5 with water on oxidized Heme a_3, while the aquo-Heme a_3 $pK_7^{'}$ is 5.9 with water on the reduced Cu_B (Table 5.2). As in the fully oxidized protein, the mutual repulsion of the two hydroxyls means that only one will be found. Thus, there is little proton uptake into the BNC on the reduction of Cu_B, but the hydroxide migrates from Cu_B to Heme a_3.

In the ORORstate, the aquo-Heme a_3 $pK_7^{'}$ shifts up to 8.4 (Table 5.2). Hence, a proton will be taken into the BNC at pH 7 coupled to the reduction of Heme a if Cu_B is already reduced. The change in $pK_7^{'}$ from OOOR to ORORstate is due to the aquo–Heme a_3 interaction with the neutral, reduced Heme a being 2.5 ΔpK units smaller than it is with the cationic, oxidized heme. After two electrons enter the BNC to form the R (OORR) intermediate, the aquo-cofactor's pK_as are greater than 7 so both remain protonated as calculated previously (66). This result agrees with experiments that show little proton release at pH 7 when CO is photolyzed off the mixed valence complex, initiating back electron transfer moving from the OORR to ORORstate (67). The pK_a obtained from the fractional site ionization in Monte Carlo sampling as a function of pH for aquo-Heme a_3 is 8.6 is in good agreement with the measured pK_a of 9 for proton release.

The pK_as of the BNC aquo-cofactors have been calculated by Siegbahn and colleagues using the hybrid density functional method (157). Since no other residues are allowed to change ionization, their pK_as are equivalent to $pK_7^{'}$ here. In the DFT study, the $pK_7^{'}$ is <–10 for the first deprotonation with 2 waters in the site, and 8 when forming 2 hydroxides in the OOOO state. In the MCCE calculations, the

.

pK_as are –9.4 and 10.7. Thus, both studies agree that there will be at least one hydroxyl in the BNC. But, the stability of the system with two hydroxides differs by 3 ΔpK units. In the DFT calculations, 10% of the oxidized BNCs would have 2 hydroxyls, while MCCE predicts none will.

The protein is treated differently in the two simulations. MCCE includes the whole protein in the model, while the DFT calculations include only the cofactors, their ligands, the propionates, and Tyr 288. The DFT simulation region is well chosen to make the net interaction of the BNC with other parts of the protein small. The missing interactions from the rest of the protein favor the hydroxyl groups by only ≈1 ΔpK unit. A more significant difference is that the DFT calculation assumes that the BNC is in a uniform medium with ε=4, while MCCE puts ε=80 into protein cavities. The high dielectric cavity stabilizes the more highly charged water-cofactor species, shifting the aquo-Heme a_3's $pK_7^{'}$ up by ≈1 pH unit, and the hydroxyl–Cu_B by ≈5 pH units. This term destabilizes the second hydroxide in the MCCE calculations. MCCE calculations were made eliminating the continuum water in the cavities and the interactions with parts of the protein outside of the DFT simulation region. Here the $pK_7^{'}$s is –15.1 for aquo-Cu_B as it loses the first proton and 8.3 for aquo-Heme a_3 forming the second hydroxyl in the BNC, significantly closer to the DFT values. In the OOOR state there is only one stable hydroxyl in either calculation. It has a pK_a of 7.6 in the DFT calculations, and 5.9 in the standard MCCE calculations. In the MCCE calculations using DFT assumptions, the extra solvation energy drops the aquo-Heme a_3 $pK_7^{'}$ to 5.2.

Ionization of Tyr I-288. Tyr I-288 is at the end of the K-channel, hydrogen-bonded to the water on Heme a_3 (Fig 5.2). Oxygen reduction chemistry is likely to involve a coupled electron (158-160) and proton (161-163) transfer from this Tyr to O_2. The Tyr could also serve as a proton acceptor when the BNC is reduced, if it is deprotonated in the fully oxidized state. The MCCE calculated pK_a is 8.7 ($pK_7^{'}$ 8.6), so it is 6% deprotonated at pH 7. The relatively high $pK_{a,sol}$ and large desolvation energy

.

destabilizes its ionization, while the BNC positive charge favors ionization (Table 5.3). The $pK_7^{'}$ is lowered by the $pK_{a,sol}$ being 8.9, 1.3 pH units lower than a standard Tyr, because of its attachment to His I-284 (161). Thus, this Tyr is neutral even in the oxidized protein, but there is only a 2.2 kcal/mol penalty for forming the anion at pH 7 in the equilibrated protein with a hydroxyl on Cu_B.

Ionization of Glu I-286. Glu I–286 is an essential residue which been proposed to shuttle chemical protons from the D channel to the BNC, and pumped protons to the outside at the appropriate steps in the reaction cycle (39-41). In the fully oxidized protein, Glu I-286 is calculated to be fully protonated with a pK_a of >10 ($pK_7^{'}$ 9.9). Despite the low $pK_{a,sol}$ of 4.8, the loss of reaction field energy destabilizes the ionized form, keeping the Glu neutral (Table 5.3). Nearby cavities leading to the D-channel (164,165), and to Heme a_3 (45), solvate the Glu to lower $pK_7^{'}$. Interactions with the protein, mostly contributed by the oxidized BNC, further stabilize Glu^-. The balance of favorable interaction with the protein and the large desolvation energy tunes the $pK_7^{'}$ so that it takes only 4 kcal/mol to deprotonate the Glu at pH 7. This puts the ionized state low enough in energy that it can serve as an intermediate in proton transfer.

Glu I-286 has been measured to have a pK_a of 9.4 in the F state (166). The F state with a water-Cu_B(II) and a ferryl Heme a_3 (Fe(IV)=O^{-2}), has the same net charge as the water-Cu_B(II) and ferric hydroxyl-Heme a_3. In the latter state, the Glu I-286 $pK_7^{'}$ would be 8.9 while it is 10 for Tyr I-288. Here Glu I-286 is 3% ionized at pH 7, and becomes 50% ionized at 10.9, titrating with a shallow pH dependence in reasonable agreement with the experimental value. The calculated Tyr pK_a is much higher than its $pK_7^{'}$ which is obtained at pH 7 where Glu I-286 is neutral. Unfavorable interactions with the Glu which is ionized first with increasing pH makes it harder to deprotonate the Tyr. Thus, the ionization of the Glu and Tyr are tightly coupled with an interaction energy of 2.6 ΔpK units. Once one of them is ionized, the pK_a of the other moves well above 10. The charge distribution in the BNC

.

determines which amino acid is ionized first. In an OOOO state with a hydroxyl-Heme a_3 (similar to the F state charge distribution), the pK_a for the Glu is lower than that of the Tyr. But in the lower energy OOOO state with hydroxyl-Cu_B, the Tyr has the lower pK_a.

Ionization of the heme propionic acids. All of the heme propionic acids are calculated to be deprotonated in the fully oxidized state (Table 5.1). The propionates on the A- and D-rings (nomenclature as in the PDB file) of Heme a_3, and on the D-ring of Heme a, are close to the BNC and have been suggested to lie on the proton pumping pathway (45,68-70). All the propionic acids are stabilized by the presence of Arg I-481 and I-482. Each D-ring propionate is within hydrogen-bonding distance of an Arg. The Heme a_3 A-ring propionate is further stabilized by the Mg^{2+} and its ligands: His I-411, Asp I-412 and Glu II–254. Even though the net charge of the Mg cluster is zero, the propionate is closer to the Mg^{2+} than to its anionic ligands (Fig 5.3). A hydrogen bond from the neutral His I-411 further stabilizes the charged state. The Heme a A-ring propionic acid is stabilized by Arg 52. The Heme a_3 propionate $pK_7^{'}$s remain below zero in all of the oxidation states, despite their increasing 2 to 5 pH units when the BNC is reduced (Table 5.2). Any changes in protonation of the propionic acids would need to be coupled to movement of the adjacent Arg, as well as to reduction of the BNC (167). There is little difference in the propionic acid $pK_7^{'}$s for OOOR and OORR states, because the second reduction of the BNC is coupled to the protonation of the hydroxide, therefore it is electroneutral with little long-range electrostatic impact (Table 5.2).

The Heme a A-ring propionic acid has a $pK_7^{'}$ near physiological pH (Table 5.2). It remains fully ionized when the BNC is reduced (OORR). But, when the BNC and Cu_A and Heme a are reduced (RRRR) its $pK_7^{'}$ is 6.1, so it is partially protonated at pH 7 (Table 5.1). However this proton binding in the fully reduced protein, is to the propionate furthest from the BNC and it is coupled to the reduction of Cu_A and Heme a, not the BNC cofactors.

.

The Heme a_3 A–ring propionic acid $pK_7^{'}$ has also been estimated by Siegbahn and colleagues, using DFT calculations (157). They found that in the fully oxidized protein it is 5.2, while MCCE finds it to be –8.0. This large difference is a result of the DFT calculations being centered at the BNC cofactors, locating the propionic acids at the edge of the simulation region about 8Å from the heme iron. Arg I-481, I-482, and the Mg^{2+} complex are not included in the calculation. These groups together stabilize the acid ionization by over 13 ΔpK units. MCCE calculations with only the groups used in the DFT calculations gives a $pK_7^{'}$ of 6.6, in reasonable agreement with the DFT analysis. In the OOOR state, DFT calculates the propionate $pK_7^{'}$ to be 11.5, while it is –4.5 in MCCE considering the whole protein, and 11.4 calculated by MCCE considering only the residues included in the DFT simulations. In the OORR state, the $pK_7^{'}$ is calculated to be 11.5 with DFT calculations, –3.9 with MCCE considering the whole protein, and 12.0 with MCCE considering the same residues as in the DFT calculation. Hence, the classical electrostatics calculations can reproduce the DFT $pK_7^{'}$s when the same region of the protein is considered. The simple classical calculation have the advantage of being able to routinely include the whole protein.

Ionization of His I-334. His I–334, a Cu_B ligand, has also been proposed to act as a residue whose protonation state is coupled to the BNC redox state. Earlier work by Wikström suggested that this His might break its bond to Cu_B and be doubly protonated in a His shuttle (71). More recently, Stuchebrukhov proposed that the His becomes a fully deprotonated imidazolate, coexisting with a hydroxyl group on Cu_B in the fully oxidized state. This His^- could then serve as a second proton acceptor, keeping the BNC reduction electroneutral (72,73). A $pK_7^{'}$ was calculated here, using a DFT calculated $pK_{a,sol}$, for deprotonation of a His ligand on an oxidized water-Cu_B of 9.0 (73). The His^- reduces the water-Cu_B complex charge, so it has a smaller desolvation penalty then the neutral His (Table 5.3). The His^--water-Cu_B is stabilized by ≈14 ΔpK units, when it is moved to a uniform medium

.

with ε=4 in both the MCCE calculations and DFT self-consistent reaction field calculations (73). Within the protein, cavities and the large unfavorable interactions with the Heme a_3 propionates destabilize His^-. The resultant $pK_7^{'}$ of 10.3 with a hydroxyl-Heme a_3 indicates that the His will remain neutral in the fully oxidized state.

In order to estimate the $pK_7^{'}$ for His deprotonation with a hydroxyl on Cu_B, a second approach is used (see methods). Here $pK_7^{'}$ is estimated starting with an isolated imidazolate $pK_{a,sol}$ of 14.4, and then including explicit non-bonded pair-wise interactions between the His and Cu_B. Despite the fact that this model is likely to over-stabilize the anionic ligand due to the strong interaction between the metal center and the ligand, the His $pK_7^{'}$ is >14 (Table 5.3). Previous calculations by Stuchebrukhov using a similar model for the His-Cu_B interaction and a continuum electrostatic based methodology similar to that found in MCCE (168-172) reported a His^- pK_a of ≈ 6 in the oxidized protein with a hydroxyl on Cu_B (72). However since the factors that contributed to their calculated pK_a were not reported, it is not clear what the critical differences are between the two calculations.

Protonation changes on residues outside the active site: At equilibrium, after reducing both cofactors in an initially oxidized protein, one hydroxide on a BNC cofactor has picked up a proton. In addition, other more distant residues change ionization state, contributing to the net proton uptake. The protonation of distant sites can be perturbed either by changes in the long-range electrostatic potential when the BNC is reduced, or by changes in local environment if conformational rearrangements are triggered by reduction of the BNC cofactors. Electroneutral reduction of the BNC cannot yield long-range electrostatic changes. In the analysis presented here, where no backbone motions are allowed, there must be a change in the net BNC charge for proton uptake to distant residues to be coupled to electron transfer. The residues that are sensitive to small shifts in the long-range electrostatic potential are partially protonated with pK_as near 7. Two clusters, one on the proton input and one on the release

.

side, make the largest contributions to the total proton uptake. On cofactor reduction, there are shifts in charge distribution within these clusters as well as changes in net cluster ionization.

The ionization states of His I–93 and Glu I–182, 5 Å apart on the proton output side of the protein, are coupled together (Fig 5.3a). At low pH (<5) one proton is bound with 13% $HisH^{+}Glu^{-}$ and 87% $His^{0}GluH^{0}$. Both microstates have a net charge of zero and similar energies, so both are occupied. The doubly ionized state is destabilized by the desolvation penalty for the two ionized groups, but stabilized by the favorable interaction between them (173). The cluster titrates with a pK_a of 6.7 in the OOOO state, producing a net charge of –0.64 at pH 7 (Fig 5.4a, Table 5.2). An electron on Cu_A shifts the pK_a up to 7.7, while the pK_a is 7.2 when Heme a is reduced. The pK_a shifts of 1.0 and 0.5 lead to average protonation changes of 0.44, or 0.27 going from OOOO to ROOO or OROO states (Table 5.1). The pK_a shifts down to 6.9 with the first reduction of the BNC. The cluster pK_a is the same in the OOOR and OORR states, since the second reduction is coupled to the protonation of the BNC hydroxyl group. And electroneutral reduction has no significant impact on residues far from the active site. When the BNC is doubly reduced, this cluster has 0.09 more protons bound than in the fully oxidized protein. Upon the anaerobic 4-electron reduction, the cluster binds 0.59 protons. In earlier calculations on *P. denitrificans* showed Lys II-191 (229 here) was involved in proton release rather than the residues shown here.

.

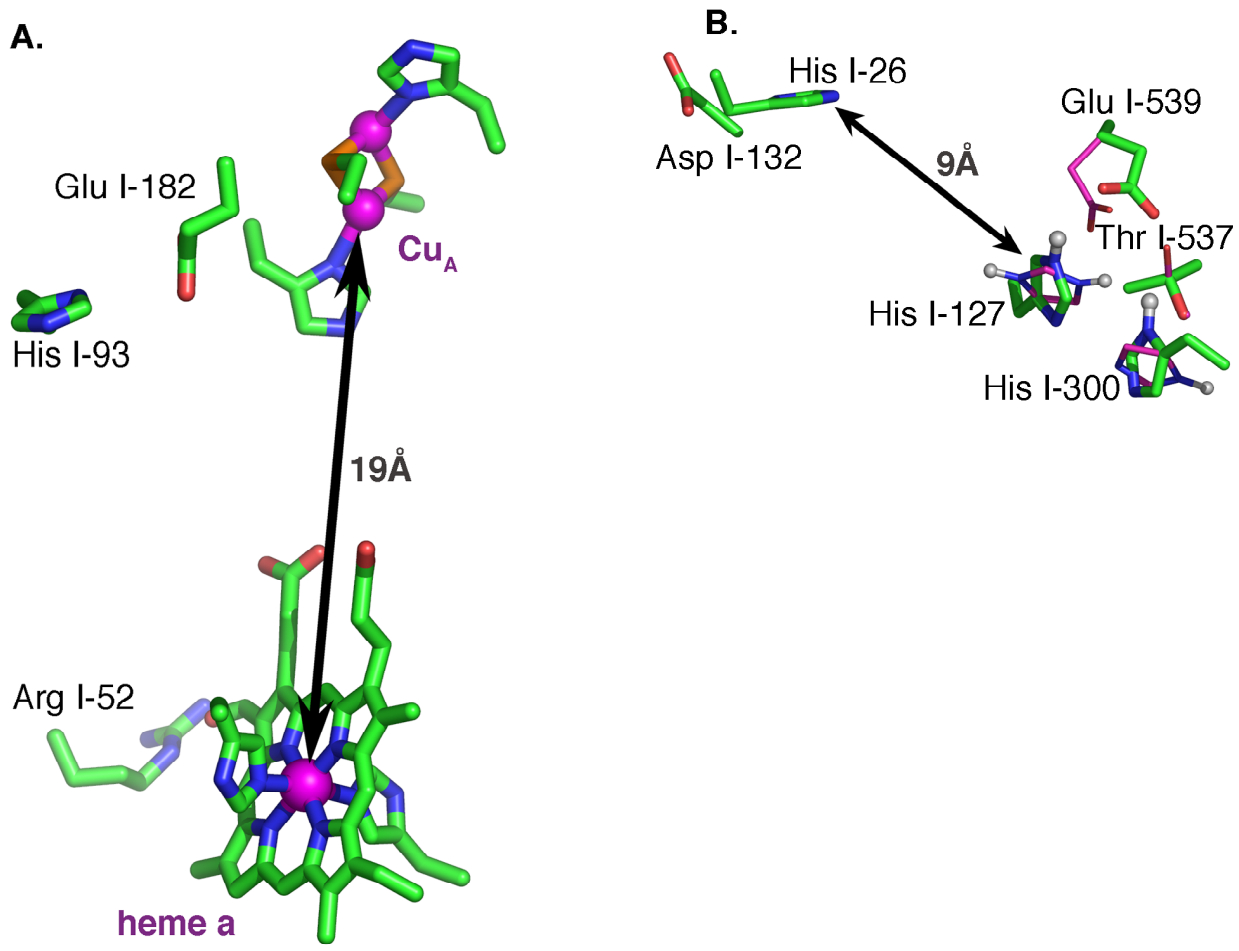

Figure 5.2 Residues around the proton release cluster and proton entry cluster

(A) Cu_A and Heme a with their ligands, Arg 52 and the proton release side cluster, His I-93 and Glu I-182.

(B) Asp I-132 and His I-26 at the D-channel entrance and the proton entry side cluster, His I-127 and Glu I-539 together with His I-300 and Thr I-537. Conformational changes in the cluster when I-127 is protonated are shown with thinner purple sticks. Protons bonded to nitrogens of His I-127 and His I-300 are shown as spheres. Glu I-539 and Asp I-132 are ≈30Å from Cu_B.

.

His I–127 and Glu I–539 form a cluster on the proton uptake side of the protein, close to Asp I–132 at the D-channel entry (Fig 5.3B). These are far from any cofactor, and more weakly coupled to cofactor redox changes than the output-side cluster. In the OOOO state, the cluster has a pK_a of 7.4, which shifts to 8.0 in the RRRR state, leading to the uptake of 0.15 protons for 4 electron reduction at pH 7.0 (Fig 5.4B). Conformational changes of Glu^- I-539, His I-300, and Thr I-537 are calculated to be coupled to the cluster ionization changes. This reorganization of the hydrogen bonding network could slow the cluster ionization changes, modifying the proton transfer to this cluster, which is ≈9 Å from the D channel entry.

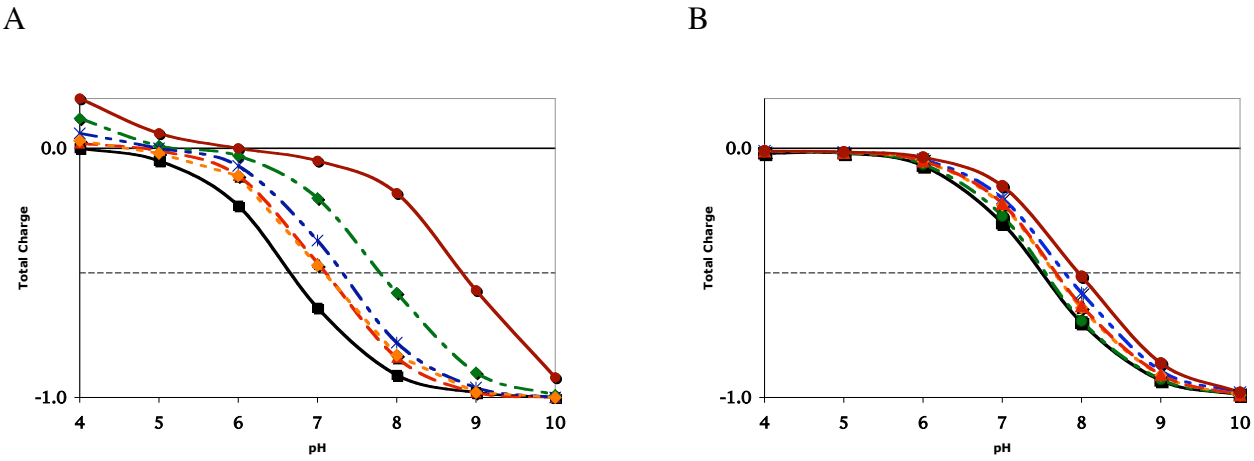

Figure 5.3. pH and redox state dependence of the net charge on clusters in cytochrome c oxidase

(A) Proton release cluster: His I-127 and Glu I-539. (B) Proton entry cluster: His I-93 and Glu I-182. (black) OOOO (6.5, 7.4), (green) ROOO (7.6, 7.4), (blue) OROO (7.2, 7.6), (orange) OOOR (6.9, 7.5), (red) OORR (6.9, 7.6) and (brown) RRRR (8.6, 8.0) states. The pK_a for the cluster 0 to –1 transition for the (release, entry) cluster is given in parenthesis. The pK_a for the ORORstate is (7.6, 7.9) (not shown).

Asp I-132 is at the D-channel entrance (174). It is always ionized with a pK_7' <1 in all oxidase redox states (Table 5.2). Since Asp I-132 is not deeply buried, it has a small desolvation penalty (Table

.

5.3). In addition, nearby backbone dipoles, and an ionized His 26 stabilize the ionized Asp. Glu II-101 is at the K-channel entrance (42,43,66). It is also kept ionized by a small desolvation penalty, and favorable interactions with the backbone and other nearby residues including His II-96, which is ≈ 30% ionized at pH 7. This residue is equivalent to Glu II-78 in *P. denitrificans*, a residue that was previously calculated to become more protonated as the BNC is reduced (66). Lys I-362 is an essential conserved residue on the K channel (44). Although, there are some suggestions that it is ionized (175), MCCE calculates it to be fully neutral with a $pK_7^{'} < 0$ in all redox states. It is deeply buried, and has few favorable interactions with neighboring residues or backbone dipoles that would stabilize the charged state. It is far enough from the BNC, that its ionization is only moderately destabilized by the cofactor positive charges.

Net proton uptake on reduction of cytochrome c oxidase. MCCE calculates that the reduction of Cu_A, in the fully oxidized enzyme is coupled to the uptake of 0.6 protons to the proton exit cluster and a few nearby groups (Table 5.1). When the electron is transferred from Cu_A to Heme a, the release side cluster looses 0.17 proton back to solution, while the uptake side cluster picks up 0.14 proton. All together, 0.5 protons are bound on reduction of Heme a relative to the fully oxidized state. The calculations show negligible additional proton uptake as the electron is transferred from Heme a to Cu_B.

There is a wide range of reported experimental values for the stoichiometry of proton uptake coupled to Cu_A and Heme a reduction. Oxidation of the fully reduced oxidase with a CO ligand to Heme a_3 (RRRR to OORR states) shows release of 0.3 (176), 0.4 (74), 0.6 (75), 0.8 (177) protons in bovine, and 0.8 protons in *P. denitrificans* (51) oxidase. However, the E_m of Cu_A and Heme a (54,177) vary by no more than 20 mV per pH unit. This is in better agreement with the smaller estimates of proton uptake. The proton uptake coupled to Cu_A and Heme a reduction in a protein, with the BNC reduced, is calculated to lead to the binding of 0.9 $H^+/2e^-$. Almost all of the change is in the exit cluster, with a

.

small amount in the proton input cluster. There are no changes in the BNC associated residues moving from OORR to RRRR states. The MCCE value of 0.5 H^+/e^- means Cu_A and Heme a would each have E_ms with a pH dependence of at least 30 mV/pH unit. The overestimation here is most probably due to the sensitivity of the His I–93 and Glu I–182 cluster to the electrostatic potential of Cu_A. This cluster binds 0.44 protons on the oxidation of Cu_A, and releases 0.28 protons as the electron is transferred to the BNC.

The second reduction, forming either the ORORor OORR mixed valence states at pH 7, leads to the binding of one proton to the hydroxyl-Heme a_3. The pK_a of the aquo-Heme a_3 is 8.3 in the ORORstate; so it is ≈90% protonated. Thus, with Cu_B reduced, the reduction of Heme a, not a_3, is calculated to be coupled to the proton binding into the BNC (Table 5.1). When the second electron is transferred from Heme a into the BNC, there are only small changes in total protonation, in agreement with the experiments photolyzing the CO-Heme a_3 bond in the mixed valence OORR state (67). Reduction of all 4 cofactors moving from OOOO to RRRR state is coupled to the uptake of 2.5 protons (Table 5.1). Most of this is accounted for by the hydroxide in the BNC, proton uptake and release side clusters, and the Heme a A-ring propionate, together with small changes on a few other distant groups. This is in agreement with the total proton uptake of 2.4 upon full reduction, which was measured in Bovine oxidase (74,75). Yet, it is smaller than a recent measurement of 3.3 H^+/4 electrons in *P. dinitrificans* (51).

The E_m of Heme a. MCCE has been used successfully to calculate the heme E_ms in different bis-His cytochromes (63). This allows the Heme a E_m to be calculated with no free parameters. The calculated E_m of 360mV, when the other cofactors are oxidized, is in agreement with the experimental values of 340 mV (54) to 430mV (76). The E_m is shifted 480 mV from the $E_{m,sol}$ of –120mV mostly by the desolvation penalty and the positive potential from the backbone amide dipoles. Although the BNC and groups such as the heme propionates, Arg I-481 and I-482, individually have a large effect on the heme

.

E_m, these roughly cancel. Therefore, all residues lower the E_m by only 70mV. This match between experimental and calculated E_m is obtained with all of the Heme propionates ionized; despite earlier suggestions by DFT calculations that the E_m would be too low (–150mV) without protonation of one of the Heme a propionates (155). One of the propionic acid charges was removed from the DFT calculations to compensate for not including the nearby Arg in the simulation region.

The Heme a E_m has been found to depend on the BNC redox state. It is lowered by 80~135mV when both Cu_B and Heme a_3 are reduced (54,77,78,178). In the OORR state, the E_m is calculated to be 260 mV, 100mV lower than in the oxidized enzyme; in good agreement with experiment. Simply adding two electrons to the cofactors, and one proton to the BNC hydroxide, lowers the E_m by 150mV. Proton uptake to other distant residues diminishes the E_m shift. If the BNC reduction were electroneutral with $2H^+/2e^-$, the heme E_m would be independent of the BNC redox state. If there are two hydroxyls in the oxidized BNC, the Heme a E_m is 260, the same as calculated in the fully reduced BNC, which is significantly lower than measured in the oxidized protein.

.

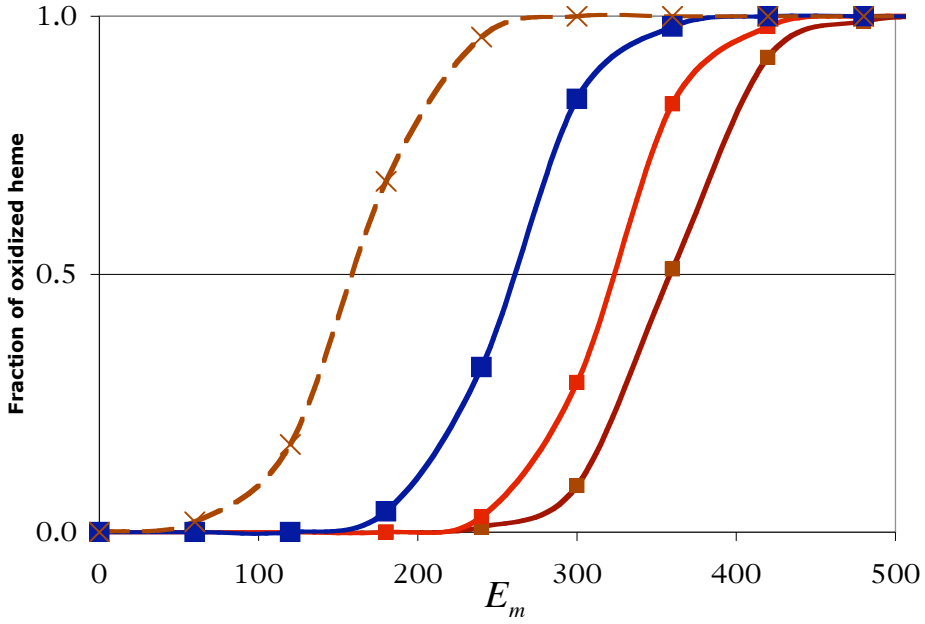

Figure 5.4. Redox titration of Heme *a*

Redox titration of Heme a with (brown) the BNC fully oxidized (E_m=359mV), (red) Heme a_3 oxidized and Cu_B reduced (E_m=322mV), (blue) the BNC doubly reduced (E_m=258), (brown dashed) the BNC fully oxidized with neutral Arg 52, as in the Arg to Met mutation (E_m=161mV).

The Heme a E_m was also calculated with Arg 52 mutated to Met. This has been observed to lower the Heme a E_m by about 260mV in *P. denitrificans* oxidase (178). In MCCE with the *Rb. sphaeroides* structure, removal of the Arg charge causes the protein to bind ≈0.5 protons. About half of this is to the Heme a A-ring propionate. The resultant Heme a E_m with the other cofactors oxidized is calculated to be lowered by 200 mV, in reasonable agreement with experiment.

.

Discussion

The number of protons bound at each stage of the electron transfer cycle in cytochrome c oxidase has been the subject of intensive study (16,35-37). The desolvation penalty for reducing the deeply buried BNC has been suggested to be the primary source of the tight coupling between electron and proton transfers needed for pumping in oxidase (52,74,179). Transferring an electron to a cofactor deep in the protein introduces a large desolvation penalty, which can be diminished by burying another charge with an opposite sign, such as a proton, nearby. Anions in the surroundings stabilize oxidation of the deeply buried cofactors, and cofactor reduction increases their pK_as to favor proton binding (Table 5.2). However, in the equilibrium calculations on the *Rb. sphaeroides* oxidase only one group, the hydroxyl on Cu_B, shifts its pK_a from below 7 to above 7 to pick up a proton when 2 electrons are transferred into the BNC. Similar conclusions were found in the MCCE analysis of the *P. denitrificans* structure, where only one proton is also taken up on formation of the R (OORR) state (66). Only one proton is needed for the O–O bond splitting chemistry of bound O_2 in the A to P_R transition (17,52,180). This proton and an electron are proposed to be donated by Tyr I-288, forming a hydroxyl on Cu_B and unprotonated Heme a_3 ferryl species ($Fe^{4+}=O^{2-}$) (158-163).

At some point in the reaction cycle following O_2 bound splitting, Heme a_3 and Cu_B each will have a hydroxyl ligand with a bound chemical proton (52). This work has shown that the stable BNC has one hydroxide and one water in the fully oxidized state, which marks the end of the oxidative phase and the beginning of the reductive phase. Thus three chemical protons need to be bound before reaching the stabilized fully oxidized state. However in the O state, a deprotonated Tyr I-288, hydroxyl on Heme a_3, or anionic His^- I-334 are only 1.6~3.5 ΔpK units higher than the equilibrated state, also allowing a second anion to be accessible in the oxidative phase. This raises the possibility that the equilibrated O state observed in MCCE calculations is that of the stabilized "resting" (180,181) enzyme, while a fully

.

oxidized protein with two deprotonated groups in the BNC is a higher energy, metastable state which occurs during the reaction cycle.

Relative energy of different BNC ionization states. In order for a group to be a proton acceptor when the BNC is reduced, it must have a pK_a below 7 in the oxidized protein and above 7 when the cofactors are reduced. Only the hydroxyl-Cu_B satisfies the criteria for a group whose protonation state is tightly coupled to the BNC reduction (Fig 5.6). All propionic acid pK_7's remain below 7 in OOOO and OORR states. The pK_7's of His I-334, aquo-Heme a_3, Glu I-286, and Tyr I-288 remain above 8. These ionization patterns are in agreement with previous calculations on *P. denitrificans* (66) and bovine (72) oxidase. Tyr I-288, with a pK_a of 8.7 (pK_7' 8.6) in the fully oxidized state, has a pK_a in the different redox states closest to what would be needed for it to be involved in proton coupled electron transfer. The Tyr is calculated to be ≈5% deprotonated in the OOOO state; so it contributes a small amount to the proton uptake coupled to the first BNC reduction. Proton uptake via the K channel, as would likely be used to deliver protons to the Tyr, has been found experimentally on the first reduction of the BNC (48,49). The relatively low Tyr pK_a relies on its $pK_{a,sol}$ being lowered 1.3 pH units by its covalent linkage to His I-284 (161). With a standard $pK_{a,sol}$ of 10.2, the cost of deprotonating the Tyr would be 2.9ΔpK units. Tyr pK_7' would be 9.9.

.

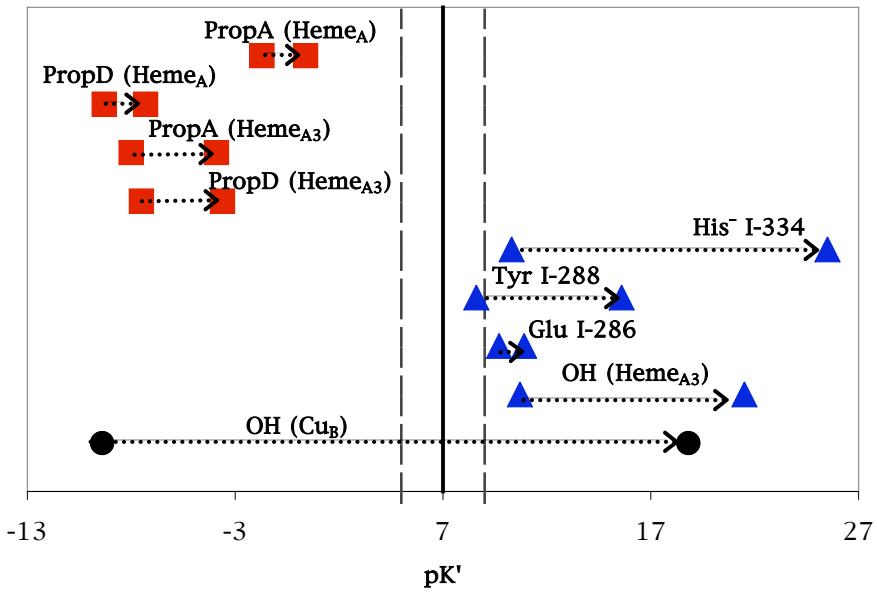

Figure 5.5. $pK_7^{'}$ **of the residues near the BNC in OOOO and OORR states**

In all cases, $pK_7^{'}$ *moves to higher values in the more reduced protein.* $pK_7^{'}$ *fixes all other residues in their lowest energy ionization state in MCCE calculations at pH 7 (Table 5.1). Thus, there is a hydroxyl on* Cu_B *in the OOOO state and no hydroxyls in the BNC in the OORR state. The 4 propionic acids are found to be ionized and Tyr I-288, Glu I-286, and His I-334 are neutral in both states.*

The energy of the different ionization states of the possible proton acceptors, including the hydroxyls on Heme a_3 and Cu_B, Tyr I-288 and His I-334, can be compared in the fully oxidized *Rb. sphaeroides* cytochrome c oxidase (Table 5.4). The highest energy state is where none of these groups are ionized, and the lowest has the hydroxyl-Cu_B as the sole anion. The hydroxyl-Heme a_3, water-Cu_B state is 4.4 ΔpK units higher. States with a single anion on the Tyr, or His, are at still higher energy. Thus, the lowest energy state has only one group that can serve as a proton acceptor when the BNC is reduced. The di-anion hydroxyl-Cu_B, ionized Tyr state is only 1.6 ΔpK units (2.2 kcal/mol) higher, so

.

this is the most accessible state with 2 proton acceptors. All other di-anion states are at significantly higher energy, even in the fully oxidized protein.

Proton binding on each reduction of the BNC has been studied experimentally. There is evidence that the stoichiometry should be 2 $H^+/2e^-$ upon reduction during the O to R transition, rather than 1.4, with only one proton bound into the BNC, as found here (Table 5.1). Using photoexcited Ruthium bispyridyl to deliver single electrons to the *P. denitrificans* oxidase BNC, electrometric measurements of wild type and Lys I-362 to Met mutants showed one electrometric phase is inhibited in the mutants, suggesting the first reduction is coupled to the proton uptake via the K-channel (48,49). Multiple excitation accumulates the doubly reduced enzyme, which is associated with a slower electrometric phase, attributed to a second proton uptake (49). Proton uptake during anaerobic reduction of the fully oxidized, CO bound oxidase also supports a model where there are two protons bound on the two electron reduction of the BNC (74).

In contrast, the good match between the calculated and measured shift in the Heme a E_m when the BNC is reduced provides evidence that there is only $1H^+/2e^-$ (54,77,78,178). If the BNC reduction is electroneutral, there is insufficient change in the long-range electrostatic potential to shift the E_m of a residue 13 Å away. In addition, the lack of proton release at pH 7 moving from the OORR to ORORstates, following flash photolysis of a bound CO (67) is in good agreement with the protonation of the BNC being coupled to Heme a reduction in the OOOR state, as found here.

Assumptions and uncertainties in the analysis. The conclusion reported here, that there is one anionic group in the fully oxidized BNC that is strongly coupled to reduction, relies on a number of assumptions. The repulsion between the cationic oxidized water-Heme a_3 and the hydroxyl-Cu_B(II) in the fully oxidized (OOOO) and mixed valence (ORORstates has been shown, by comparison between Gaussian98 and Coulomb's law, to be overestimated in a continuum electrostatics treatment (119). Using

.

the uncorrected interactions, the pK_a of the aquo-Heme a_3 is low enough that both of the BNC cofactors bind hydroxyl in the fully oxidized protein. The same correction for interactions between nearby charges is needed to obtain benchmark pK_as in agreement with experiment for aquo-hemes in two other proteins (119). In addition, a model system with a Cu_B analogue inserted into myoglobin also shows a surprisingly small change in the water-Heme E_m when a Cu or a Zn is added (182).

A very simple metal centered charge distribution is used for the cofactors in the BNC. The long-range interactions are only affected by the net charge, but local interactions are sensitive to the partial charges on individual atoms. Using Gaussian98 charges for the hydroxyl-3His-Cu_B complex moves some of the positive charge from the copper to the His ligands, further from the hydroxyl on Heme a_3. This shifts the aquo-Heme a_3 pK_a up by ≈2 pH units, indicating that both of the hydroxyls are even less likely to coexist in the oxidized BNC with a more distributed charge model.

Protein cavities are filled with high dielectric continuum water, instead of explicit water in the calculations. Buried waters have been suggested to be important for oxidase activity (45-47,165,183,184). Earlier studies of bacteriorhodopsin have shown that the calculated pK_as are similar with either continuum or explicit water, even for active site residues (114). In addition, if explicit waters are not placed carefully, necessary hydrogen bonds cannot be formed to compensate for the desolvation penalty incurred by filling in the cavities with explicit waters, introducing additional errors.

Kinetic vs thermodynamic control of ionization states. This analysis provides a view of the equilibrium protonation states throughout the protein in the *Rb. sphaeroides* oxidase crystal structure. The MCCE calculated $pK_7^{'}$ provides the free energy needed to change the ionization of a residue, providing clues to which higher energy ionization states can be kinetically accessible. Glu I-286 and Tyr I-288 are both near the BNC at the end of the proton uptake channels. They have $pK_7^{'}$s between 9 and 10 in the fully oxidized state, as well as in the P_R and F states in the oxidative phase. Hence, they are

.

protonated at equilibrium, but they need only 2.5-4 kcal/mol for them to give up their proton. In contrast, the heme propionates on the output side of the protein have pK_7's below 0, making it very difficult for them to be protonated without changes in the protein structure. These calculations may overestimate the cost of deprotonating the propionates. Arg motion could diminish the cost of acid protonation, raising the pK_a.

Calculations on the bacteriorhodopsin proton pump shows that small changes in the structure can lead to directional protonation shifts between residues, while each residue remains in equilibrium in the given structure. These calculations are carried out with high-resolution structures trapped in different states of the bacteriorhodopsin reaction cycle. MCCE (114) and MEAD (152) calculations with structures differing by <1Å RMSD show changes in the active site ionization, as predicted by experiment (21). The agreement between the kinetic measurements and equilibrium calculations shows that the protonation states remain in equilibrium with the structure; at least for processes occurring on the slow microsecond time scale. The challenge for cytochrome c oxidase is to find the appropriate conformational changes required to stabilize the proton transfer intermediates. Glu I-286 has been found changing conformation during an MD simulation (185). If this Glu can move towards the positively charged BNC, the ionized form would be stabilized, lowering the pK_a. The propionic acids are hydrogen bonded to a pair of Args. If propionic acids and the Arg move apart to break the ion pair, one of the propionic acids could be protonated (167). His I-334 is hydrogen bonded to two propionic acids. If the propionic acids move away from the His, the unfavorable interaction between negative imadazolate and the propionic acids would be reduced; perhaps, allowing His I–334 deprotonation. Analysis of protonation states of residues equilibrated around these modified positions are ongoing.

.

A. Net charge	OOOO	ROOO	OROO	OOOR	OROR	OORR	RRRR
His I-93	0.05	0.17	0.12	0.08	0.12	0.08	0.22
Glu I-182	-0.69	-0.37	-0.49	-0.56	-0.47	-0.54	-0.27
Release side Cluster	**-0.64**	**-0.20**	**-0.37**	**-0.48**	**-0.35**	**-0.46**	**-0.05**
PropA (Heme a)	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-0.79
aquo-Heme a_3	0.00	0.00	0.00	-0.87	-0.09	0.00	0.00
aquo-Cu_B	-1.00	-1.00	-1.00	0.00	0.00	0.00	0.00
Glu I-286	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tyr I-288	-0.06	-0.04	0.00	0.00	0.00	0.00	0.00
BNC Cluster	**-1.06**	**-1.04**	**-1.00**	**-0.87**	**-0.09**	**0.00**	**0.00**
His I-127	0.31	0.29	0.48	0.37	0.44	0.44	0.50
Glu I-539	-0.62	-0.62	-0.67	-0.59	-0.64	-0.66	-0.66
Uptake side Cluster	**-0.31**	**-0.33**	**-0.19**	**-0.22**	**-0.20**	**-0.22**	**-0.16**
Whole protein	**-4.30**	**-4.70**	**-4.80**	**-4.90**	**-4.80**	**-4.90**	**-5.80**
B. Proton uptake relative to OOOO state		ROOO	OROO	OOOR	OROR	OORR	RRRR
Release side Cluster	**-**	**0.44**	**0.27**	**0.16**	**0.29**	**0.18**	**0.59**
BNC Cluster	**-**	**0.02**	**0.06**	**0.19**	**0.97**	**1.06**	**1.06**
Uptake side Cluster	**-**	**-0.02**	**0.12**	**0.09**	**0.11**	**0.09**	**0.15**
Whole Protein	**-**	**0.60**	**0.50**	**0.40**	**1.50**	**1.40**	**2.50**

.

C. Step-wise proton uptake		OOOO → ROOO	ROOO → OROO	OROO → OOOR	OOOR → OROR	OROR → OORR	OORR → RRRR
Release side Cluster	**-**	**0.44**	**-0.17**	**-0.11**	**0.13**	**-0.11**	**0.41**
BNC Cluster	**-**	**0.02**	**0.04**	**0.13**	**0.78**	**0.09**	**0.00**
Uptake side Cluster	**-**	**-0.02**	**0.14**	**-0.03**	**0.02**	**-0.02**	**0.06**
Whole Protein	**-**	**0.60**	**-0.10**	**-0.10**	**1.10**	**-0.10**	**1.10**

Table 5.1 Ionization states of key residues.

MCCE calculated ionization of key residues of Rb. sphaeroides cytochrome c oxidase in selected redox states; The whole protein ionization includes charges on groups not explicitly noted here. Redox states of the 4 cofactors are represented by the four-letter key ordered: Cu_A, Heme a, Heme a_3 and Cu_B (O: oxidized, R: reduced). The A and D propionic acids are identified by their connection to the A and D rings in the PDB file. The IUPAC-IUB designation is D and C respectively.

.

	OOOO	ROOO	OROO	OOOR	OROR	OORR	RRRR
BNC cluster							
aquo-Heme a_3	10.7	11.1	12.4	5.9	8.4	21.5	22.6
aquo-Cu_B	-9.4	-9.1	-7.8	26.2	16.8	18.8	20.7
Glu I-286	9.7	9.5	11.3	11.2	10.6	10.9	12.8
Tyr I-288	8.6	8.8	9.8	15.7	11.3	15.6	16.9
* His I-334 (a)	10.3	10.9	12.0	N.A.	N.A.	N.A.	N.A.
* *His I-334 (b)*	*15.8*	*16.4*	*17.5*	*22.9*	*31.8*	*25.5*	*27.7*
PropA Heme a	-1.7	0.0	2.1	-0.1	3.5	0.4	6.1
PropD Heme a	-9.3	-7.7	-6.4	-7.5	-4.8	-7.3	-3.7
PropA Heme a_3	-8.0	-7.2	-6.1	-4.5	-2.3	-3.9	-1.6
PropD Heme a_3	-7.5	-6.6	-5.4	-3.8	-1.6	-3.6	-0.8
Release side cluster							
His I-93	6.5	7.6	7.2	6.9	7.6	6.9	8.5
Glu I-182	6.3	7.3	6.9	6.7	7.3	6.7	8.1
Uptake Side cluster							
His I-127	7.8	7.7	8.0	8.0	8.3	8.0	8.4
Glu I-539	6.9	7.0	7.1	7.0	7.2	7.1	7.2

.

D and K channel residues							
Asp I-132	0.5	0.6	0.6	0.6	0.6	0.6	0.6
Lys I-362	-2.7	-2.7	-2.9	-2.1	-2.8	-2.0	-1.2
Glu II-101	2.1	2.1	2.2	2.3	2.4	2.4	2.4

Table 5.2 $pK_7^{'}$ **of residues implicated in oxidase activity.**

*$pK_7^{'}$ gives the effective pK_a calculated with the protein ionization and conformation states fixed at pH 7 providing the free energy of site changing ionization at pH 7 (Eqn. 5,6). $pK_7^{'}$ depends on the ionization states of other residues. $\Delta G_{res,7}^{mfe}$ is calculated with key residues ionization states given in Table 5.1. * $pK_7^{'}$ for deprotonation of His I-334 to form the anion is calculated in two ways, as described in Materials and Methods. (a) The His has a $pK_{a,sol}$ of 9.0 in the oxidized water-Cu_B complex (40), $pK_7^{'}$ is calculated with a hydroxyl on Heme a_3 and water on Cu_B in OOOO, ROOO, OROO states. This value is used for OOOO state in Table 5.4 and Fig 5.6. (b in italics) The His pK_a is calculated, starting with a $pK_{a,sol}$ of 14.4 and explicit, non-bonded pairwise interactions with Cu_B and its ligands. The equilibrium state here has a hydroxyl on Cu_B in OOOO state. This value is used for OORR state in Fig 5.6.*

.

	crg reactant	crg product	$pK_{a,sol}$	$\Delta\Delta G_{rxn}$	$\Delta\Delta G_{pol}$	$\Delta G^{mfe}_{res,7}$					
						OOOO		OORR		RRRR	
						BNC	RES	BNC	RES	BNC	RES
ferric aquo-Heme a_3	0	+1	9.6	3.9	0.4	0.2	-5.7	-	-	-	-
ferrous aquo-Heme a_3	-1	0	10.9	-10.4	0.4	-	-	4.0	-4.6	4.0	-5.7
cupric aquo-Cu_B	+1	+2	9.4	14.7	0.4	10.0	-6.3	-	-	-	-
cuprous aquo-Cu_B	0	+1	9.4	-4.7	0.4	-	-	1.1	-6.2	1.1	-8.1
Glu I-286	0	-1	4.8	6.8	0.8	-4.3	1.6	-2.7	1.2	-2.7	3.1
Tyr I-288	0	-1	8.9	10.4	0.3	-12.4	1.3	-4.8	0.7	-4.8	2.1
His I-334*	+2	+1	9.0	-8.5	-2.1	2.2	9.7	N.A.	N.A.	N.A.	N.A.
	0	*-1*	*14.4*	*14.0*	*-2.1*	*-20.0*	*9.7*	*-9.0*	*8.6*	*-9.0*	*10.4*
PropA Heme a	0	-1	4.9	10.4	-5.5	-3.8	-7.7	-1.8	-7.5	-1.8	-1.8
PropD Heme a	0	-1	4.9	9.8	-7.3	-4.3	-12.3	-2.1	-12.5	-2.2	-8.8
PropA Heme a_3	0	-1	4.9	7.1	-1.3	-10.4	-8.3	-5.1	-9.6	-5.1	-7.3
PropD Heme a_3	0	-1	4.9	6.5	-3.2	-9.9	-5.8	-5.2	-6.7	-5.2	-3.8
His I-93	0	+1	6.5	2.9	-0.7	0.7	-2.9	0.4	-2.9	0.4	-4.5
Glu I-182	0	-1	4.8	4.8	-2.7	-0.7	0.2	-0.4	0.2	-0.4	1.6

.

His I-127	0	+1	6.5	4.8	4.4	0.7	-11.3	0.3	-11.1	0.3	-11.5
Glu I-539	0	-1	4.8	6.7	-4.1	-0.5	0.0	-0.2	-0.1	-0.2	0.1
Asp I-132	0	+1	4.8	3.3	-2.9	-0.3	-4.3	-0.2	-4.4	-0.2	-4.4
Lys I-362	0	+1	10.8	10.9	-0.7	2.6	1.2	1.6	1.0	1.7	0.2
Glu II-101	0	-1	4.8	2.7	-1.2	-0.6	-3.5	-0.3	-3.6	-0.3	-3.5

Table 5.3 Energy terms used to calculate $pK_7^{'}$.

*The net charge of the reactant and product states is given. The reference reactant state is neutral for: Prop, Glu, Lys and His, has a neutral His I-334 and a hydroxyl on Heme a_3 and Cu_B; $pK_{a,sol}$: pK_a of isolated group in aqueous solution; $\Delta\Delta G_{rxn}$: double difference of desolvation (reaction field) energy $[product - reactant]_{in\ protein} - [product - reactant]_{in\ water}$. This is generally a positive, unfavorable term when the product charge is larger than that of the reactant. ΔG_{pol} and $\Delta G_{res,7}^{mfe}$: Differences in electrostatic and non-electrostatic interactions of product and reactant states with the backbone (ΔG_{pol}), with the binuclear center cofactors, and their water or hydroxyl ligands (BNC $\Delta G_{res,7}^{mfe}$), and with the protein sidechains and other cofactors (RES $\Delta G_{res,7}^{mfe}$). $\Delta G_{res,7}^{mfe}$ uses a mean field energy interaction with the ionization states found by MCCE sampling at pH 7 (see Table 5.1 for key residues). Thus, in the OOOO state there is a water on Heme a_3 and a hydroxyl on Cu_B. The hydroxyl moves to Heme a_3 in the OOOR state. There are no hydroxyls in the BNC in any other redox states. For example, the $\Delta G_{res,7}^{mfe}$ BNC term for ferric aquo-Heme a_3 in the OOOO state is the (ferric-water-Heme a_3) – (ferric-hydroxyl-Heme a_3) interaction with the cupric hydroxyl-Cu_B, the MCCE calculated equilibrium state for Cu_B. The interaction of the ferric-water heme and each of the Cu_B states is reduced as shown in supplementary information (see methods and (53)). * $pK_7^{'}$ of His I-334 deprotonation is calculated in two ways, as*

.

described in Materials and Methods. The top entry uses a Cu(II)-water-His complex. This does not allow calculations with a hydroxyl on Cu(II) or any Cu(I) states. The BNC interactions are with hydroxyl-Heme a_3 and its ligands only. In the bottom entry, the BNC $\Delta G^{mfe}_{res,7}$ term includes the non-bonded interactions of the Cu, the other 2 Cu_B His, and its aquo ligands, as well as Heme a_3 with His I-334.

.

State energy pH 7 (ΔpK units)	OH^- Heme a_3	OH^- Cu_B	His I–334*	Tyr I–288	
Neutral state					
16.4	0	0	0	0	
single anion					site $pK_7^{'}$ †
0	**0**	**-1**	**0**	**0**	**-9.4**
4.4	-1	0	0	0	-5.0
10.7	0	0	0	-1	1.3
11.7	0	0	-1	0	2.3
di-anionic states					
1.6	0	-1	0	-1	
3.3	-1	0	-1	0	
3.7	-1	-1	0	0	
10.1	0	0	-1	-1	

Table 5.4 Relative energy of different protonation states of BNC cluster residues in the fully oxidized BNC.

*Energies are relative to the state with single OH^- on Cu_B (in bold). Ionization of all other key residues are given in Table 5.1. *Using model for His in complex with water-Cu_B with $pK_{a,sol}$ of 9.0 (40). † $pK_7^{'}$ for groups assigned a –1 charge. This is different from that given in Table 5.2 since those are calculated*

.

with ionization given in Table 5.1 while here the BNC residues are in the explicitly defined, non-equilibrium ionization states. 1 ΔpK unit is 1.36 kcal/mol.

.

6. Calculations of aquo-heme pK_a

The ferric aquo-heme pK_as were calculated with MCCE in myoglobins from sperm whale and *Aplysia*, hemoglobin I, heme oxygenase 1, horseradish peroxidase and cytochrome c oxidase and compared to measured values. These pK_as span 3.3 pH units from 7.6 in heme oxygenase 1 to 10.9 in the peroxidase. Two sets of calculations are presented. In the first set, standard Poisson-Boltzmann continuum electrostatics (CE) calculations are used for all pair-wise interactions between the aquo-heme and the surrounding protein residues. In the second, an additional correction derived from DFT calculations is applied. Calculated pK_as are compared to the experimental values. Calculations of the heme E_ms are discussed below.

The pK_as calculated using the standard CE based pairwise interactions in MCCE (63) do not provide consistent agreement with the experimental data (Fig 6.1). Calculated pK_as for the myoglobins and heme oxygenase are within 1 pH unit of the measured values. But errors as large as 6 pH units are found in horseradish peroxidase and cytochrome c oxidase. In these proteins a nearby positive charge in the heme distal pocket lowers the calculated pK_a, while the myoglobins have no charges in similar positions. As described in the Methods section CE and DFT calculations provide very similar values for pairwise interactions between a charge and a neutral hydroxyl-heme. For the cationic water-heme, both methods also yield similar results when the charge is on the heme edge. However when the charge is above the heme plane near the iron, the interactions derived with a CE analysis are significantly larger than those obtained with DFT calculations. The closer the charge is to the heme plane, the larger the discrepancy (see Fig 2.2). The DFT corrections are applied to the pairwise electrostatic interaction of the water-heme with Arg99 in hemoglobin I, Arg38 in horseradish peroxidase, and Cu_B in cytochrome c oxidase to recalculate the pK_as. Now calculated pK_as in the six proteins agree with the measured values

.

within 1 pH unit. Analysis of the MCCE calculations shows how interactions and conformational changes shift the $pK_{a,sol}$ of 9.6 to different in situ values in these proteins.

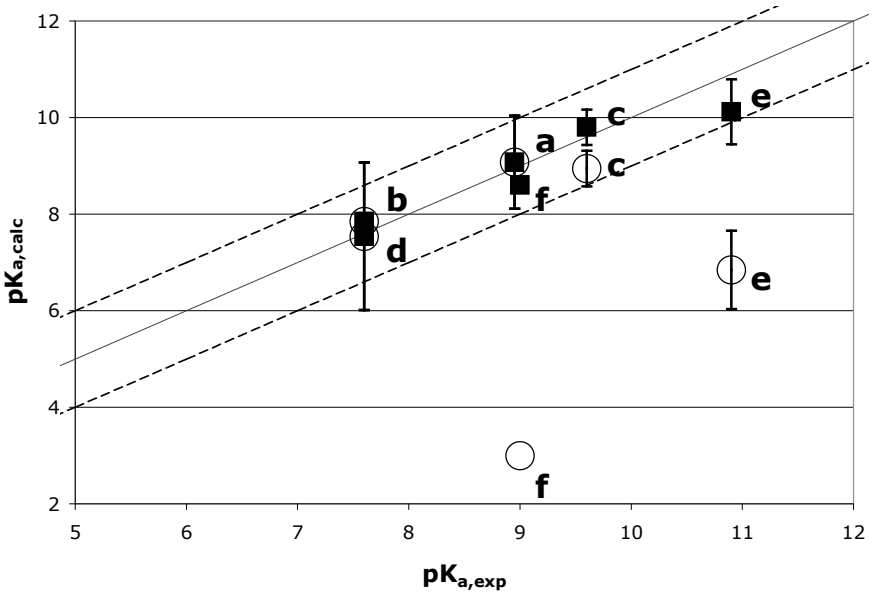

Figure 6.1 Comparison of calculated and experimental pK_as

(■) with and (○) without the DFT correction for (a) sperm whale myoglobin, (b) Aplysia myoglobin, (c) hemoglobin I, (d) heme oxygenase, (e) horseradish peroxidase and (f) cytochrome c oxidase. Ideal line of slope 1 passing through the origin and lines with differences between experimental and calculated of ±1 pH unit (dashed lines) are shown. References for experimental values in Table 6.1.

.

Table 6.1 Experimental and calculated pK_as and E_ms and energy terms contributing to pK_a and E_m shifts in the protein

a. pK_a of ferric aquo-heme.

	Exp pK_a	MCCE pK_a	$\Delta\Delta G_{rxn}$ (ΔpK unit)	ΔG_{pol} (ΔpK unit)	ΔG_{res} (ΔpK unit)
Sperm Whale Myoglobin (7)	8.95 (186)	9.1±1.0	2.07±0.29	0.91±0.14	-2.45±0.83
Aplysia Myoglobin (2)	7.6 (186)	7.7±1.0	1.70±0.34	0.88±0.23	-0.75±0.49
Hemoglobin I (4)	9.6(187)	9.8±0.3	1.30±0.18	1.02±0.09	-2.57±0.24
Heme oxygenase 1 (5)	7.6 (188,189)	7.5±1.5	0.16±0.37	3.34±0.69	-1.44±1.10
Horseradish Peroxidase (10)	10.9 (186)	10.1±0.7	1.58±0.39	-0.61±0.07	-1.49±0.66
Cyt c Oxidase (1)	9.0(67)	8.6	3.9	0.4	-3.3

.

b. E_m of aquo-heme at pH 7.

	Exp E_m (mV)	MCCE E_m (mV)	$\Delta\Delta G_{rxn}$ (ΔpK unit)	ΔG_{pol} (ΔpK unit)	ΔG_{res} (ΔpK unit)
Sperm Whale Myoglobin (7)	50 (190)	25±33	4.50±0.56	0.98±0.15	-2.97±0.66
Aplysia Myoglobin (2)	125 (190)	126±64	4.25±1.24	0.89±0.11	-0.89±0.24
Hemoglobin I (4)	103(191)	83±15	4.33±0.17	0.85±0.12	-1.68±0.09
Heme oxygenase 1 (5)	-	37±13	2.47±0.36	1.84±0.44	-1.61±0.30
Horseradish Peroxidase (10)	-250(192,193)	-215±27	3.80±0.51	-0.62±0.06	-4.79±0.49
Cyt c Oxidase[a] (1)	-	-11	5.77	0.77	-5.70

The value in parenthesis is the number of PDB files analyzed. Negative free energy terms favor the cationic ferric water-heme raising the aquo-heme pK_a and favor the oxidized heme lowering the E_m. The ferric aquo-heme $pK_{a,sol}$ is 9.6 (126,127). The b-type water-hemes have an $E_{m,sol}$ of –140mV, while the a-type heme in cytochrome c oxidase has a $E_{m,sol}$ of –60mV. Hydroxyl-hemes have $E_{m,sol}$s 64 mV lower than the water-hemes (124,125). [a] *The pK_a and E_m of the aquo-Heme a_3 is calculated with Cu_A oxidized and Heme a and Cu_B reduced to mimic the experimental pK_a measurements in the CO photolyzed mixed-valence complex (67). 1ΔpK unit = 58 meV = 1.36 kcal/mol. A systematic shift of 0.74ΔpK unit (1 kcal/mol) is applied to the energy of ferric hydroxyl-heme and –0.5ΔpK unit (–0.67 kcal/mol) to ferrous water-heme in deriving the in situ E_ms and pK_as using Eqn. 2-4.*

.

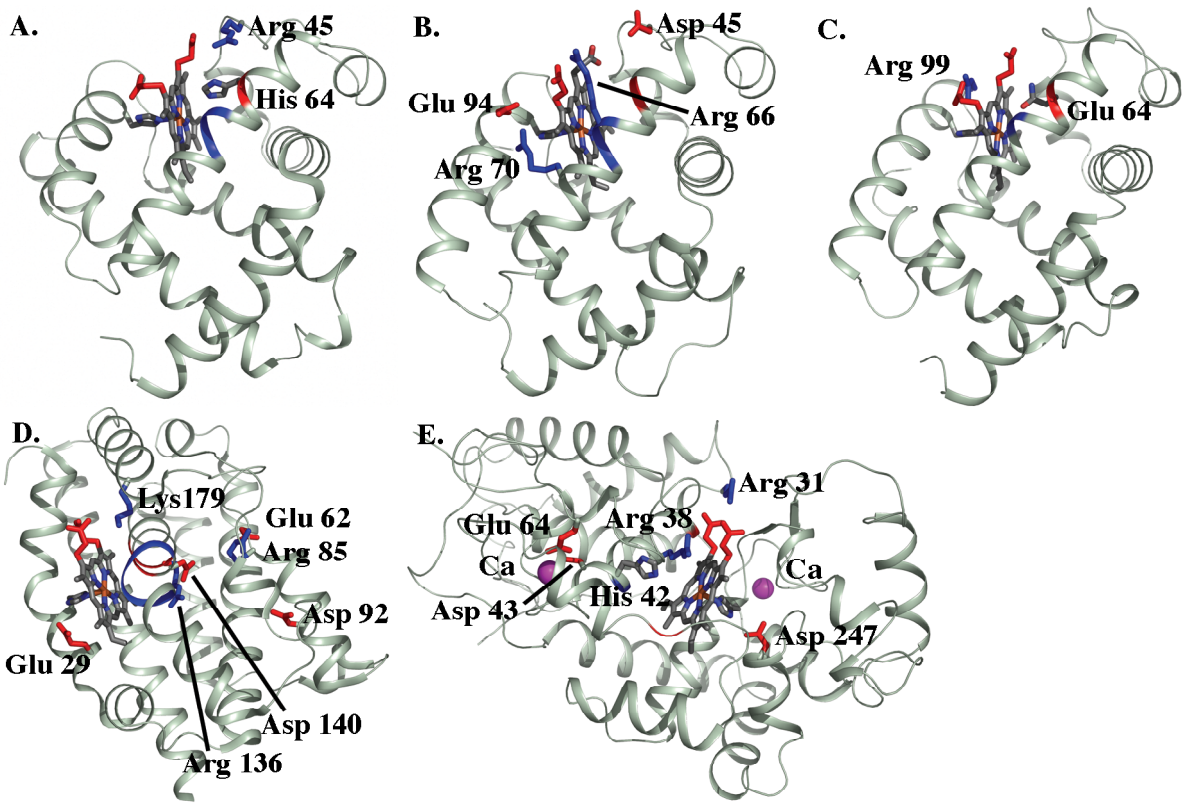

Figure 6.2. Structures of sperm whale myoglobin, Aplysia myoglobin, hemoglobin I, heme oxygenase, and horseradish peroxidase

(A) sperm whale myoglobin, (B) Aplysia myoglobin, (C) hemoglobin I, (D) heme oxygenase, (E) horseradish peroxidase. Residues and backbone segments that interact with the cationic ferric aquo-heme by more than 0.5 ΔpK units (0.68 kcal/mol) are shown.

.

Sperm whale myoglobin. Sperm whale myoglobin has a folded 6-helix core (Fig 6.2). The pK_a of the ferric aquo-heme is 8.95 (186), little different from the $pK_{a,sol}$ of 9.6. MCCE calculations with 7 different crystal structures give an average aquo-heme pK_a of 9.1±1.0, in good agreement with experiment. This heme is close to the protein surface. However, $pK_a \approx pK_{a,sol}$ not because the heme does not interact with the protein, but rather because a number of effects cancel (Table 6.1). The water-heme, with a net +1 charge, has 7 ΔpK units (9.52 kcal/mol) less solvation energy in the protein than it does in water, but the hydroxyl-heme, which has a large dipole moment, has a 5ΔpK units desolvation energy. Thus, the resultant desolvation penalty ($\Delta\Delta G_{rxn}$) favors the water-heme complex raising the pK_a by about 2 pH units. The backbone amide dipoles tend to raise the electrostatic potential within all proteins shifting residue pK_as (59,139) and cytochrome E_ms (63). Here they favor the hydroxyl-heme, lowering the pK_a by 1 pH unit. The E helix, which contains the His 64 in the distal pocket is the largest contributor.

There are 4 Arg, 14 Glu, 7 Asp and 19 Lys in the protein and all of them are >95% ionized between pH 7 and 9. However, most of these charges have little effect on the aquo-heme pK_a because they are solvent exposed. Arg 45, which is hydrogen-bonded to the D-ring propionic acid does shift the pK_a down by 0.6 pK unit (propionate A and D designated as assigned in the PDB structure files). The two heme propionic acids are both fully ionized and together shift the pK_a up by 2.3 units. As these are all near the heme edge, the CE interaction agrees with the DFT values (Fig 2.2). His 64 in the distal heme pocket near the aquo-ligand undergoes a conformational change with the change in the aquo-heme protonation state (Fig 6.2). The His is always neutral, but the δ proton tautomer stabilizes the positive water-heme by 3~4 ΔpK units, while the protonated ε tautomer stabilizes the hydroxyl-heme by 1.6~2.5 ΔpK units. In different PDB structures the His tautomer is restricted by the rigid backbone to different extents, yielding the dependence of the calculated pK_a on the starting structure. Arg 45 not only lowers

.

the aquo-heme pK_a by direct electrostatic interaction, it also lowers the pK_a indirectly by favoring the ε His 64 conformer.

Aplysia Myoglobin. Myoglobin from the sea hare, *Aplysia limacine* has a folded 6-helix structure (Fig 6.2). It is calculated to have a ferric aquo-heme pK_a of 7.7, 2 pK units lower than that of sperm whale myoglobin, in agreement with the experimental value of 7.6 (186). The *Aplysia* and sperm whale proteins have the same protein folds and are 27% identical. The analog of myoglobin His 64 is Val 64 in the *Aplysia* distal pocket, providing a bigger cavity, so the heme is slightly better solvated shifting the pK_a up. However, the biggest difference between the two proteins is in ΔG_{res}. The loss of His 64 and protonation of the ring D propionic acid lowers the aquo-heme pK_a. The propionate pK_a is raised to 8.6 by the nearby Asp 45, while in myoglobin Arg 45 lowers the propionate pK_a to 4.6.

Hemoglobin I. Monomeric clam hemoglobin I also has a folded 6-helix structure (Fig 6.2). The aquo-heme pK_a is calculated to be 9.8 and measured to be 9.6 (187), close to that of sperm whale myoglobin. The hemoglobin I and myoglobin have similar structures with 23% identity. A larger cavity in the heme distal pocket leads to a smaller desolvation energy. As His 64 does in myoglobin, Gln 64 changes conformation keeping its oxygen pointing towards the water-heme and nitrogen pointing towards the hydroxyl-heme. Arg 99, on the proximal side of the heme shifts the pK_a down by 1.5 pH units, while the two propionic acids, both of which are ionized, shift the pK_a up by 2.2 pH units.

Heme oxygenase 1. Rat heme oxygenase 1 is a multi-helical bundle with two 3-helix motifs (Fig 6.2). The aquo–heme pK_a is calculated to be 7.5, in agreement with the experimental value of 7.6 (188,189). In this protein heme is an exchangeable substrate and is significantly more exposed than in the other proteins considered here and so has a very small $\Delta\Delta G_{rxn}$. The large positive potential from the backbone lowers the pK_a by 3.3 pH units. The distal pocket helix is kinked at Gly 143 and 144 so the backbone amides of Ser 142, Gly 143 and 144 all point toward the heme iron, stabilizing the hydroxyl-

.

heme. Arg 136 and Asp 140 are in a salt bridge near the porphyrin edge. While each interacts with the aquo-heme by over 3ΔpK units, their combined impact is small.

Horseradish peroxidase. The multihelical peroxidase uses heme to reduce hydrogen peroxide to water. The aquo-heme pK_a is calculated to be 10.1 in reasonable agreement with the measured 10.9 (186). The aquo-heme has a similar reaction field energy loss as in the globins. Unlike the globins, interactions with the backbone amides near the metal center are small and somewhat negative raising the pK_a by 0.6 pH units. Interactions with other residues raise the pK_a by 1.5 ΔpK units. His 42 and Asp 43 raise the pK_a by 6 pH units while Arg 38 shifts the pK_a down by 4.5 pH units (assuming a 60% DFT correction). His 42 sits in the distal pocket in a position similar to His 42 in myoglobin. However this His is fixed in the δ tautomer by Arg 38, Glu 64 and Asn 70, stabilizing the protonated water-heme.

Cytochrome c oxidase. Heme a_3 in cytochrome c oxidase is deeply buried in the transmembrane 12 helix subunit I of the protein. In the mixed-valence state, with Heme a and Cu_B reduced, the ferric aquo-heme pK_a has been measured to be 9 (67). Using the full CE interaction between Cu_B and the aquo-heme the calculated pK_a is 3. The DFT correction reduces the interaction from 6.4 to 3.2 ΔpK units. Since the aquo-heme titration now occurs at higher pH, the protein net charge is more negative then when the titration is at pH 3. This raises the pK_a further to 8.6, in good agreement with experiment (132). The cytochrome c oxidase heme is more deeply buried than in the other proteins, losing more reaction field energy, which shifts the pK_a down. Within the membrane embedded protein, pair-wise electrostatic interactions with other residues are quite long-range. However, the net interaction of the fully ionized propionic acids and the rest of the protein shift the pK_a up by only 2 pH units.

E_m calculations. The aquo-heme E_ms were calculated in each protein at pH 7. E_ms previously measured in 4 of the proteins range from –250mV in peroxidase to 125mV in *Aplysia* myoglobin (190)

.

(Table 6.1b). MCCE calculations agree with the experimental values with errors of ≤35 mV. MCCE allows protonation of the aquo-heme to be coupled to heme reduction (63). The relevant $E_{m,sol}$ for water-heme of –140mV (124,125) and for hydroxyl-heme of –204mV (124) is included in the microstate energy (Eqn. 1) so the pH dependence of the E_m is properly treated in the redox titration. This degree of freedom could be important for *Aplysia* myoglobin and heme oxygenase, which have pK_as near 7. However since their ferric pK_as are above 7 and ferrous pK_as are higher, the aquo-heme ligand remains predominately water at pH 7.

Electrostatic interactions that favor the ferric heme with its net charge of +1 over the neutral ferrous heme shift the E_m down. The same electrostatic interactions favor the cationic ferric water-heme over the neutral hydroxyl-heme raising the pK_a. While a higher pK_a is correlated with a lower E_m the connection is by no means exact (Fig 6.3). In most proteins, the $\Delta\Delta G_{protein}$ for the redox reaction is larger than for the protonation reaction so the E_m is shifted from $E_{m,sol}$ by more than the pK_a is shifted from $pK_{a,sol}$ (Fig 6.3, Eqn. 2-3). One difference is the desolvation energy is on average 2.4±0.3ΔpK units larger for the redox reaction than for the protonation reaction (Table 6.1). Because of its significant dipole the solvated neutral ferric hydroxyl-heme has a solvation energy that is not much smaller than the positively charged ferric water-heme. Thus, both reactant and product of the protonation reaction have similar $\Delta\Delta G_{rxn}$. The desolvation penalty, which always favors the neutral species, shifts the pK_a down by at most 2pH units. In contrast when compared to the neutral, ferrous water-heme, the cationic, ferric water-heme has much larger interactions with water. The impact of the desolvation energy on the redox reaction when the heme is buried in the protein is thus much larger.

Hemoglobin I has a pK_a which is only 0.2 pH units higher than the $pK_{a,,sol}$, while the E_m is 223 mV more positive than $E_{m,sol}$ ($\Delta\Delta G_{protein}$=3.8 ΔpK units). The difference in desolvation energy for the two reactions contributes 3 ΔpK units. In addition, Gln 64 in the distal pocket changes conformation in

.

the aquo-heme pK_a titration but does not in the E_m titration. The rigid Gln 64 raises the water-heme E_m by 1 ΔpK unit (58 meV) more than it shifts the ferric aquo-heme pK_a down. The reorientation of the Gln is part of the protein dielectric response that is captured by the explicit motions available in MCCE (103,134). The Gln motion coupled to aquo-heme protonation thus diminishes the effective pairwise interaction.

In horseradish peroxidase, $\Delta\Delta G_{protein}$ shifts the E_m down more than it shifts the pK_a up. Besides the 2.2 ΔpK unit difference in the desolvation energy, ΔG_{res} favors heme reduction 3.3 ΔpK units more than water deprotonation. Although both ferrous water-heme and ferric hydroxyl-heme have a net charge 0, they interact differently with nearby residues. The hydroxyl-heme has a large dipole with its negatively charged hydroxide pointing toward the distal pocket. Arg 38 in the distal pocket and Asp 247 on the opposite side form a dipole stabilizing the hydroxyl-heme. The difference in interactions with hydroxyl-heme and water-heme is relatively small leading to a small ΔG_{res}. On the other hand, the neutral ferrous water-heme is not stabilized by the electrostatic dipole generated by these two amino acids while interactions of the ferric water-heme with nearby negative charges shift the E_m down.

As in horseradish peroxidase, the favorable interaction of Cu_B with the hydroxyl-heme in cytochrome c oxidase makes the contribution of ΔG_{res} to the in situ pK_a smaller than found for in situ E_m. The E_m of Heme a_3 is predicted to be –10mV in cytochrome c oxidase when Cu_A is oxidized and Heme A and Cu_B reduced.

There is no experimental data available for the E_m in heme oxygenase 1. These calculations predict an E_m of 40mV. A run of 4 backbone dipoles lowers the pK_a more than the E_m. All of the other proteins studied here have similar values of ΔG_{pol} for the redox and protonation reactions.

.

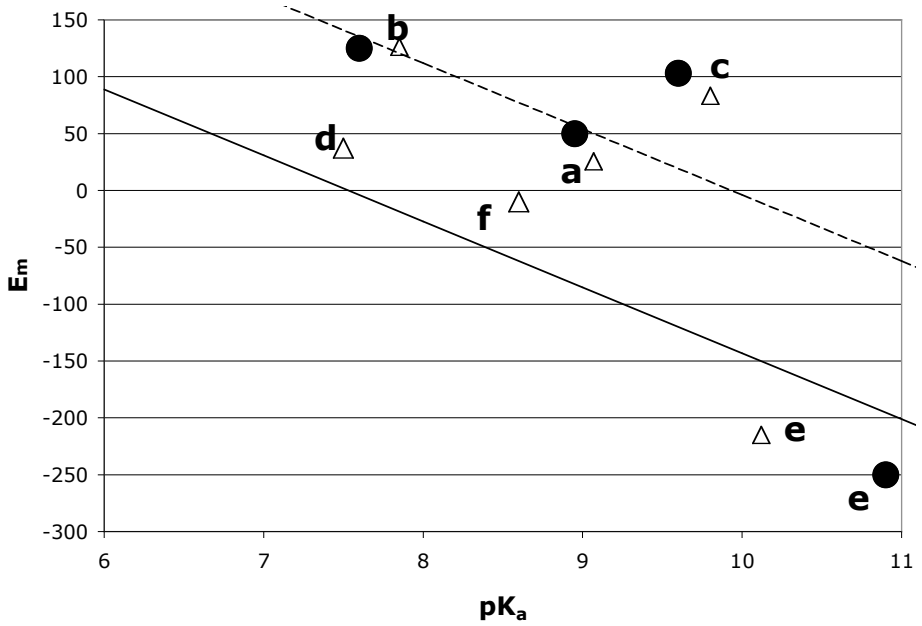

Figure 6.3. E_m versus pK_a at pH 7

(a) sperm whale myoglobin, (b) Aplysia myoglobin, (c) hemoglobin I, (d) heme oxygenase, (e) horseradish peroxidase and (f) cytochrome c oxidase. (△) Calculated and (●) available experimental values. Solid line has the same $\Delta\Delta G_{protein}$ for the in situ E_m and pK_a (Eqn. 2 and 3). Dashed line if $\Delta\Delta G_{protein}$ is 2.4 ΔpK units larger for E_m than pK_a; 2.4ΔpK units is the average difference in the $\Delta\Delta G_{rxn}$ for the two types of the reactions (Table 6.1).

Conclusions

The pK_as of ferric aquo-heme at pH 7 in sperm whale myoglobin, *Aplysia* myoblogin, hemoglobin I, heme oxygenase 1, horseradish peroxidase and cytochrome c oxidase were calculated with MCCE. The calculations reveal a breakdown in the classic continuum electrostatics analysis of pairwise interactions. Large errors in calculated pK_as were found in horseradish peroxidase and cytochrome c oxidase. DFT calculations suggest that CE calculations over-estimate pairwise

.

interactions between ferric water-heme and a positive charge above the heme plane. This could be due to significant out-of-plane charge density, which would be ignored in the atom-centered charges used in the CE analysis. Alternatively, these external charges could polarize ferric hydroxyl-heme significantly more than the amount accounted for by the ε=4 used in the MCCE calculation. Using a correction that brings the DFT and CE interactions into agreement, the calculated pK_as in six proteins are within 1 pH unit of the experimental values. The calculated E_ms using the correction derived from DFT calculations are within 35mV of the experimental values.

Proteins with higher ferric aquo-heme pK_as have lower E_ms since both arise from the protein stabilizing a positively charged heme. However, the proteins shift the free energy of the redox reaction ($\Delta\Delta G_{protein}$) by an amount that would shift a pK_a by >6 pH units (375mV), while the pK_as span only 3.3 pH units (Eqn. 2-3). The difference in $\Delta\Delta G_{protein}$ shows there is a larger effective dielectric response for the protonation than the redox reactions. In sperm whale myoglobin and hemoglobin I protonation is coupled to conformational changes while the redox reaction is not. These conformation changes allow a residue in the distal pocket to stabilize both protonation states reducing the energy difference between them. In contrast, when the group is rigid as it is for the redox reaction the pair-wise interactions are larger shifting the E_m more. This difference in flexibility for the two reaction typs found in MCCE is supported by the good match between calculation and experiment for both pK_as and E_ms.

.

Appendix

MCCE subroutines

The latest version of Multi-Conformation Continuum Electrostatics (MCCE2) (104) added improved rotatmer making and optimization to the original version of MCCE (103,112). The basic scheme of MCCE2 is described in ref. (104). The detailed algorithms of the new subroutines are described here.

IPECE (Implementing Protein Environment for Continuum Electrostatics)

IPECE is a subroutine used to place a slab of neutral atoms to provide a low dielectric membrane and to add water molecules into the protein cavities (114). There are three main parts in this subroutine: 1. The position of the membrane slab is determined and neutral atoms are added to mimic the low dielectric membrane. 2. a grid based probe to determine the protein occupied region, protein exterior space and protein cavities. 3. Cavities in the protein are probed and water or ions added into the cavities.

The added membrane slab is made of a box of neutral atoms. The slab position is optimized to bury the fewest surface ionizable residues. Here is the scheme used in the program:

1. Read in pdb file and the surface accessible area of each atom. The surface accessible area (SA) is pre-calculated using standalone SURFV program (194), and the values stored in acc.atm file.
2. All the terminal atoms of ionizable residues, Asp OD1 and OD2, Glu OE1 and OE2, Arg NH1 and NH2, and Lys NZ together with their SA are saved in a list.
3. The protein is translated so that the center of all atoms (the average position) is at the origin. This position is stored as p_1.
4. The membrane slab defined between $z = \pm\frac{1}{2}a$, where a is the thickness of the membrane.

.

5. A score number is calculated by the weighted summation of all SA from the list made in step 2. $s = \sum(\gamma_i \times SA_i)$, γ is 0 if the atom is not within the membrane slab ($z < -\frac{1}{2}a$ or $z > \frac{1}{2}a$), and for those atom within the membrane slab, γ is 1 for each terminal oxygen or nitrogen of Asp, Glu and Arg and 2 for Lys in the slab.
6. Starting form p_1 position, the protein is rotated along the x-axis by 90° (p_2) to calculate the score s again. Restore to p_1 position, and then rotated along the y-axis by 90° (p_3) to calculate the score s again. The score s from p_1, p_2, and p_3 are compared and the position with the minimum s is the starting position for the next step.
7. Then the protein is randomly shifted along the z-axis by 0~1Å, or rotate around a random axis on x-y plane by 0~15°, and the score s calculated again. If the score is lower than from the previous position, then this shift or rotation is accepted. If the score gets higher, then the chance of accepting this shift/rotation is $e^{-0.2\times s}$. A random number is generated to compared to $e^{-0.2\times s}$ to decide whether this step is accepted.

   A fully surface exposed ionizable residue contributes 60~70Å^2 to the score. Assuming that it costs 12kT to move an exposed ionizable residue into the membrane region, this leads to the coefficient 0.2kT/Å^2 used in calculating the acceptance rate. In a real case, it usually costs more than 12kT to bury a charge, however the purpose here is to sample enough positions to find the lowest scored position, therefore the energy barrier is lowered to allow the protein move more freely.
8. Step 7 is repeated until no new lowest score can be found in 20,000 success steps.
9. The protein is then moved back to the lowest score position.
10. The protein is probed. Neutral atoms are added to the grid points that are marked as X.

Assigning all probes (Fig App. 1):

1. A grid space covers the whole protein. By default, the grids are separated by 1Å.
2. For each protein atom, the nearest grid point is marked as P (protein).
3. A probe list is generated starting from the corners. All eight corners are marked as X (protein exterior) as the first generation of the probes.
4. The next generation of the probes is generated from all the neighbor grid points of the previous generation that are neither P or X. They are then marked as X.
5. Step 4 is repeated until no new probes can be found in a new generation.
6. Now, those grid points that are not marked as P or X are not accessible for protein exterior, and they are marked as C (cavities).

.

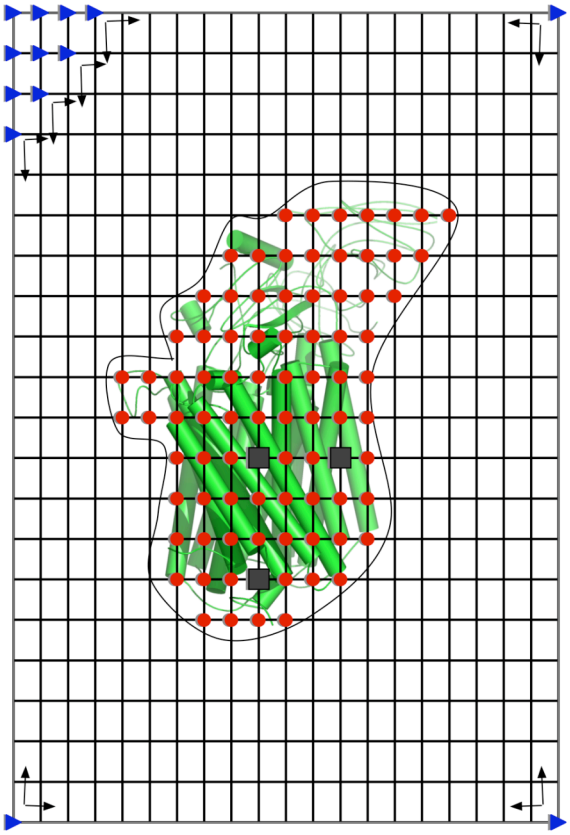

Figure App. 1 Schematic illustration of the probe procedure

The probes start from the corners of the grid box to fill the entire protein exterior space. (●) marked as P (protein), (▸) marked as X (exterior) and (■) marked as C (cavity).

Probe and membrane placing procedures are modified for the pore-forming proteins on the outer membrane. These proteins, such as porin (Fig App. 2), have a large pore through the transmembrane region (195,196). On the protein-membrane interface, the protein is hydrophobic, however inside the protein, there are many ionizable residues facing the transmembrane pore to conduct ions and substrates. In the probe procedure, now the probe list starts from the corner of the membrane layer (step 3, Fig App. 2). And the propagation of probes stays within the membrane layer. Therefore the probes cannot reach the pore if the protein is oriented properly relative to the membrane layer.

.

The calculation of the score *s* is also modified. For ionizable residues within the membrane slab, if it is more than more than 2Å away from a grid point that is marked X, then the γ is 0 so that residues in the pore are not counted into the score. And in the membrane placing procedure, each time the protein is moved, the protein is probed again before the score is calculated.

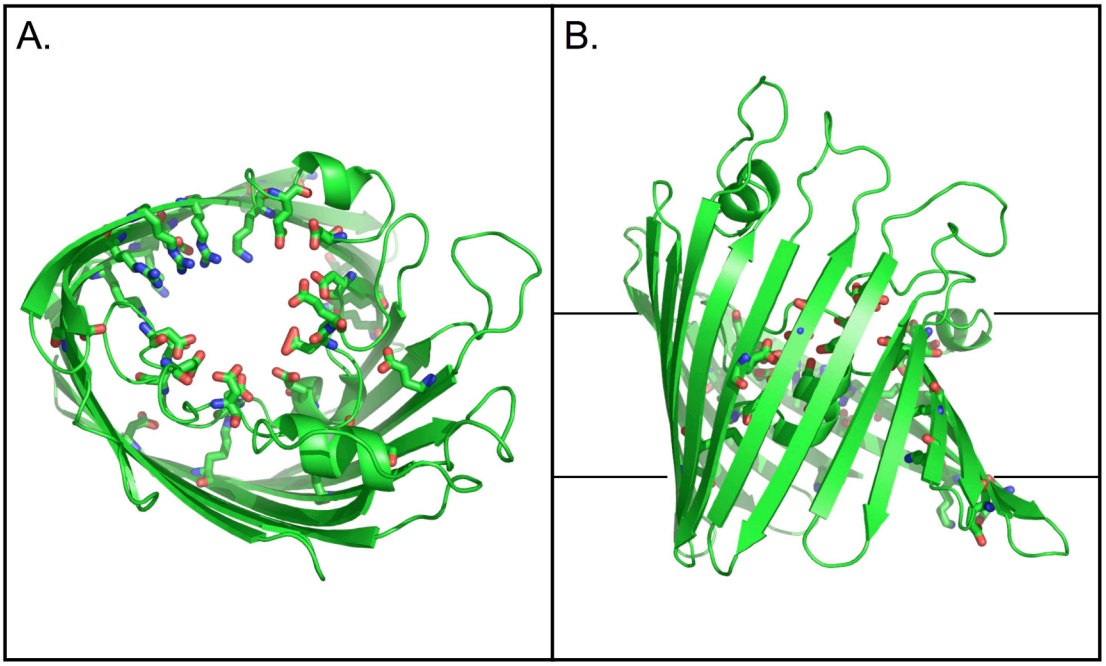

Figure App. 2 Structure of a porin

Structure of a porin (PDB: 2POR) shown in cartoon, ionizable residues in the transmembrane region are shown in sticks. (A) the trans-membrane pore is filled with ionizable residues. (B) the interface between porin and membrane is hydrophobic. In the probe procedure, the probes stay in the membrane layer.

Add water. Water or other small molecules can be added into the protein by IPECE. After the protein is probed, the grid points marked as C are not directly accessible from the solution through a corridor of the probe radius (by default is 2.2Å). Water oxygen or single atom molecules are then added

.

to these grid points. Hydrogen positions of water are generated in MCCE standard hydrogen placement subroutine.

.

REPACK

Rotamers are placed on the protein by rotating around each rotatable bond by a defined step (by default is 60º) to fill the entire phase space (104). Those with overlaps within a sidechain or between the rotamer and the rigid backbone are eliminated. The rotamers left at this stage can be further pruned to eliminate those that are likely to clash to other sidechains. The protein is repacked to find a low energy state a number of times (5000 by default). The scheme for each repack cycle is:

1. The protein starts with a microstate where each residue is occupied in a random rotamer.
2. A list is generated with all the residues, and then shuffled to create a random order.
3. Each residue is checked with the order from step 2. The energy of each rotamers is calculated, where torsion energy, Lennard-Jones energy within conformer itself and Lennard-Jones energy with the rest of the protein at current configuration are considered. All the rotamers that have energies within a threshold (5 kcal/mol by default) to the lowest are defined as low energy rotamers. If the current occupied conformer is among the low energy rotamers, no flip is needed. Otherwise, this residue changes to a rotamer that is randomly selected from the low energy ones.
4. Step 2-3 are repeated until no flip is made through the whole protein. A low energy configuration is generated at this point. Each rotamer has a counter, which increases by one if the rotamer is a low energy rotamer at this configuration.

After a number of cycles, those rotamers with low chance of being a low energy rotamer (less than 5% of cycles being counted) are eliminated.

.

RELAXATION

The repacked process is carried out with only heavy atoms. Those rotamers having low Lennard-Jones interactions may have overlapping hydrogens. A subroutine is used to carry out small energy minimization steps and relax the overlapping hydrogens.

1. In each cycle, all the conformer pairs are checked. If the heavy atom Lennard-Jones interaction is small (<5Kcal/mol by default), and the full atom Lennard-Jones interaction is large (>2Kcal/mol), then this pair is subject to energy minimization.
2. The sidechain and the backbone atoms of the selected conformer pair are included in the minimization process and the rest of the protein is not considered. Each process is composed of many iterations (50 by default). In each iteration, all backbone atoms are fixed and sidechain atoms are moved under the force field raised by the conformer pair. The displacement is $\vec{s} = \frac{1}{2}\frac{\vec{F}}{m}t^2$, where t is 1fs by default; m is mass atom; and $\vec{F}$ is combined force given by torsion energy, Lennard-Jones interaction, electrostatic interaction, and additional user defined constraints. Torsion and Lennard-Jones energies are calculated with standard AMBER force field (197). Electrostatic interactions are calculated using Coulomb's law with $\varepsilon = 1$, using the same charge assignment as later in Delphi calculations. Two constraints are added to keep the sidechains from moving too far from the initial configuration. First an energy penalty is applied if an atom is moved over a distance (1Å by default) away from the initial position that is recorded before going into the RELAXATION subroutine. The energy penalty can be calculated by $E = \frac{1}{2}k\left(\left|\vec{x} - \vec{x}_0\right| - d\right)^2$, where k is the sprint constant (10 kcal/mol/Å^2), $\vec{x}$ is the current position, $\vec{x}_0$ is the initial position, d is the distance within which no penalty is applied. The second constraint is applied to the torsion energy. Since the subroutine is mainly used to relax

.

the hydrogen clash, at the beginning of the process, the force on hydrogen can be large while the rest of the conformers are barely moved. Terminal hydrogen can be pushed into a different torsion minimum to relieve the clash and the desired relaxation between heavy atoms is not achieved. Therefore all torsion energies are increased (by 20 times by default) at the beginning of a minimization step. Then linearly scaled to standard value halfway through the minimization. The second half of the minimization process uses standard torsion energy. No bond-length or bond-angle energy is applied. SHAKE algorithm (198) is used to constrain all bond-length and angle to the initial values.

The original conformers from the crystal structure have duplicated conformers used in minimization step, and the original copy is not moved in this step. No other extra conformers are generated in this step. After several cycles (10 by default), the final position of each conformer is kept.

.

Monte Carlo

The Monte Carlo procedure is the final step of MCCE, which samples the phase space of the protein to achieve a Boltzmann distribution of all conformers. The scheme for a standard Monte Carlo procedure is:

1. The protein is initialized from a microstate where each residue is occupied in a randomly selected conformer. The system energy is calculated by adding up all the self energy and pair-wise energy (Chapter 2).
2. In each iteration, a new configuration is generated by randomly select a residue, and flip to a random conformer that was not occupied. The change of system energy ΔE is calculated. If system energy decreases (ΔE<0), then the new configuration is accepted. If system energy increases (ΔE>0), then the new configuration has a chance, $e^{-\frac{\Delta E}{kT}}$, of being accepted. A random number (ξ) between 0 and 1 is generated, and compared to $e^{-\frac{\Delta E}{kT}}$. If $\xi < e^{-\frac{\Delta E}{kT}}$, then the new configuration is accepted, otherwise the system stays in the old configuration. The counter of each occupied conformer increases by 1 at the end of the iteration.
3. The beginning of the simulation is for the system to reach low energy states from a random configuration, called equilibration period. After a number of iterations ($500 \times N_{conf}$ by default, where N_{conf} is the total number of conformers), all the counters are reset to zero. Following is the data collection period. At the end of this period, occupancy of each conformer ($iconf$) is calculated as $\frac{Couter_{iconf}}{N_{iteration}}$, where $N_{iteration}$ is the number of iterations in this period ($2000 \times N_{conf}$ by default).

.

Several techniques have been incorporated into MCCE to achieve fast convergence for large proteins. In the regions where residues have multiple ionization states and positions, and they are coupled together due to the large interactions, the energy surface is complicated. Sometimes, these coupled residues need to change conformation together to move between low energy states, while changing only one residue brings the system to a higher energy level which is hard to be accepted in a Monte Carlo simulation. For example, an Asp and Lys pair is buried inside the protein with little interaction to other residues, but strong interaction between Asp^- and Lys^+. Then Asp^-Lys^+ and Asp^0Lys^0 states have low energies as they have zero total charge, while Asp^-Lys^0 and Asp^0Lys^+ states have high energies as the energy for burying charge is not compensated by favorable interactions. Flipping between Asp^-Lys^+ and Asp^0Lys^0 states is not directly achieved in a standard MCCE flip. The convergence is hard to achieve without the enough sampling of both states, as the system can be easily trapped in one. Multi-flip technique (199) is implemented into MCCE to bypass the local high energy barrier (114).

1. For each residue (*ires*), a list (BIG LIST) is created to include all the residues that have at least one large conformer-conformer interaction with *ires* (>5 kcal/mol by default).
2. Now in each iteration, the number of residue being flipped (n_Flip) is randomly chosen between 1 and a pre-defined $N\_Flip_{\max}$ (3 by default). n_Flip is calculated after the first residue to be flipped (*kres*) is selected, and the number of residue in the BIG LIST of *kres* (N_BIG_{kres}) is checked. $N\_Flip_{\max}$ is modified to $N\_BIG_{kres}+1$ if $N\_BIG_{kres}+1<N\_Flip_{\max}$.
3. The new configuration is generated by flipping conformers of residue *kres* and (n_Flip-1) residues randomly selected from the BIG LIST of *kres*.

.

In a large protein, sometimes there are a few regions that have complicated local energy surface and hard to converge. But the rest of the protein can achieve convergence quickly. In a standard Monte Carlo simulation, the coupled residues are not explored enough. A modified procedure is used here to focus the sampling more on these regions.

1. A weighted residue list is created, initialized with one copy of each residue. In each iteration, $kres$ is now selected from this list.
2. The whole Monte Carlo is divided into small periods of iterations (10,000 by default). At the end of each period, the average occupancy of each conformer from this period is recorded. A standard deviation ($stdev\_occ_{iconf}$) is calculated for each conformer using recorded occupancies. The maximum value is recorded for each residue ($stdev\_occ_{ires,\max}$). The larger this value is, the harder it is to converge for this residue. The weighted list is then updated with $\left(1+1000\times stdev\_occ_{ires,\max}\right)$ copies of each residue $ires$.

Therefore fore each residue, the less converged it is, the more chance is has to be explore. Since this list only bias the chosen residue, not trial Metroplis flip of a conformer, the bias on final distribution is negligible. The calculations with and without this technique differ by less than 5% for all conformer occupancies.

.

References:

1. **Mitchell, P.** 1961. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. Nature **191**:144-148.
2. **Kalckar, H. M.** 1937. Phosphorylation in kidney tissues. Enzymologia **2**:47-53.
3. **Kalckar, H. M.** 1991. 50 years of biological research--from oxidative phosphorylation to energy requiring transport regulation. Annu Rev Biochem **60**:1-37.
4. **Lipmann, F.** 1939. Coupling between pyruvic acid dehydrogenation and adenylic acid phosphorylation. Nature **143**:281-283.
5. **Lipmann, F.** 1941. Metabolic generation and utilization of phosphate bond energy. Advances Enzymol **I**:99-105.
6. **Hendriks, J., U. Gohlke, and M. Saraste**. 1998. From NO to OO: nitric oxide and dioxygen in bacterial respiration. J Bioenerg Biomembr **30**:15-24.
7. **Sagan, C., W. R. Thompson, R. Carlson, D. Gurnett, and C. Hord**. 1993. A search for life on Earth from the Galileo spacecraft. Nature **365**:715-721.
8. **Woese, C. R.** 1987. Bacterial evolution. Microbiol Rev **51**:221-271.
9. **Stoeckenius, W., R. H. Lozier, and R. A. Bogomolni**. 1979. Bacteriorhodopsin and the purple membrane of halobacteria. Biochimica et biophysica acta **505**:215-278.
10. **Lanyi, J. K.** 2000. Bacteriorhodopsin. Biochim Biophys Acta **1460**:1-3.
11. **Luecke, H., B. Schobert, H. T. Richter, J. P. Cartailler, and J. K. Lanyi**. 1999. Structure of bacteriorhodopsin at 1.55 A resolution. J Mol Biol **291**:899-911.
12. **Lanyi, J. K.** 2004. Bacteriorhodopsin. Annu Rev Physiol **66**:665-688.
13. **Ferreira, K. N., T. M. Iverson, K. Maghlaoui, J. Barber, and S. Iwata**. 2004. Architecture of the photosynthetic oxygen-evolving center. Science **303**:1831-1838.

.

14. **Kok, B., B. Forbush, and M. McGloin**. 1970. Cooperation of charges in photosynthetic O2 evolution-I. A linear four step mechanism. Photochem Photobiol **11**:457-475.

15. **Tommos, C., and G. T. Babcock**. 2000. Proton and hydrogen currents in photosynthetic water oxidation. Biochim. Biophys. Acta **1458**:199-219.

16. **Ferguson-Miller, S., and G. T. Babcock**. 1996. Heme/copper terminal oxidases. Chem. Rev. **96**:2889-2907.

17. **Brzezinski, P., and G. Larsson**. 2003. Redox-driven proton pumping by heme-copper oxidases. Biochim. Biophys. Acta **1605**:1-13.

18. **Lanyi, J. K.** 1993. Proton translocation mechanism and energetics in the light-driven pump bacteriorhodopsin. Biochim. Biophys. Acta **1183**:241-261.

19. **Lanyi, J. K.** 1997. Mechanism of ion transport across membranes. Bacteriorhodopsin as a prototype for proton pumps. J Biol Chem **272**:31209-31212.

20. **Haupts, U., J. Tittor, and D. Oesterhelt**. 1999. Closing in on bacteriorhodopsin: progress in understanding the molecule. Annu Rev Biophys Biomol Struct **28**:367-399.

21. **Balashov, S. P.** 2000. Protonation reactions and their coupling in bacteriorhodopsin. Biochim. Biophys. Acta **1460**:75-94.

22. **Luecke, H.** 2000. Atomic resolution structures of bacteriorhodopsin photocycle intermediates: the role of descrete water molecules in the function of this light-driven ion pump. Biochim Biophys Acta **1460**:133-156.

23. **Rummel, G., A. Hardmeyer, C. Widmer, M. L. Chiu, P. Nollert, K. P. Locher, I. I. Pedruzzi, E. M. Landau, and J. P. Rosenbusch**. 1998. Lipidic Cubic Phases: New Matrices for the Three-Dimensional Crystallization of Membrane Proteins. J Struct Biol **121**:82-91.

.

24. **Luecke, H., B. Schobert, H. T. Richter, J. P. Cartailler, and J. K. Lanyi**. 1999. Structural changes in bacteriorhodopsin during ion transport at 2 Ångstrom resolution. Science **286**:255-261.

25. **Lanyi, J., and B. Schobert**. 2002. Crystallographic structure of the retinal and the protein after deprotonation of the Schiff base: the switch in the bacteriorhodopsin photocycle. J Mol Biol **321**:727-737.

26. **Schobert, B., J. Cupp-Vickery, V. Hornak, S. Smith, and J. Lanyi**. 2002. Crystallographic structure of the K intermediate of bacteriorhodopsin: conservation of free energy after photoisomerization of the retinal. J Mol Biol **321**:715-726.

27. **Edman, K., P. Nollert, A. Royant, H. Belrhali, E. Pebay-Peyroula, J. Hajdu, R. Neutze, and E. M. Landau**. 1999. High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. Nature **401**:822-826.

28. **Royant, A., K. Edman, T. Ursby, E. Pebay-Peyroula, E. M. Landau, and R. Neutze**. 2000. Helix deformation is coupled to vectorial proton transport in the photocycle of bacteriorhodopsin. Nature **406**:645-648.

29. **Facciotti, M. T., S. Rouhani, F. T. Burkard, F. M. Betancourt, K. H. Downing, R. B. Rose, G. McDermott, and R. M. Glaeser**. 2001. Structure of an early intermediate in the M-state phase of the bacteriorhodopsin photocycle. Biophys J **81**:3442-3455.

30. **Luecke, H., B. Schobert, J. P. Cartailler, H. T. Richter, A. Rosengarth, R. Needleman, and J. K. Lanyi**. 2000. Coupling photoisomerization of retinal to directional transport in bacteriorhodopsin. J Mol Biol **300**:1237-1255.

.

31. **Rouhani, S., J. P. Cartailler, M. T. Facciotti, P. Walian, R. Needleman, J. K. Lanyi, R. M. Glaeser, and H. Luecke**. 2001. Crystal structure of the D85S mutant of bacteriorhodopsin: model of an O-like photocycle intermediate. J Mol Biol **313**:615-628.

32. **Pettei, M. J., A. P. Yudd, K. Nakanishi, R. Henselman, and W. Stoeckenius**. 1977. Identification of retinal isomers isolated from bacteriorhodopsin. Biochemistry **16**:1955-1959.

33. **Braiman, M., and R. Mathies**. 1980. Resonance Raman evidence for an all-trans to 13-cis isomerization in the proton-pumping cycle of bacteriorhodopsin. Biochemistry **19**:5421-5428.

34. **Scharnagl, C., and S. F. Fischer**. 1996. Conformational flexibility of arginine-82 as source for heterogeneous and pH- dependent kinetics of the primary proton transfer step in the bacteriorhodopsin photocycle: An electrostatic model. Chemical Physics **212**:231-246.

35. **Brzezinski, P.** 2004. Redox-driven membrane-bound proton pumps. Trends Biochem. Sci. **29**:380-387.

36. **Gennis, R. B.** 2003. Some recent contributions of FTIR difference spectroscopy to the study of cytochrome oxidas. FEBS Lett. **555**:2-7.

37. **Wikstrom, M.** 2004. Cytochrome c oxidase: 25 years of the elusive proton pump. Biochim. Biophys. Acta **1655**:241-247.

38. **Babcock, G. T., and M. Wikström**. 1992. Oxygen activation and the conservation of energy in cell respirations. Nature **356**:301-308.

39. **Junemann, S., B. Meunier, N. Fisher, and P. R. Rich**. 1999. Effects of mutation of the conserved glutamic acid-286 in subunit I of cytochrome c oxidase from Rhodobacter sphaeroides. Biochemistry **38**:5248-5255.

.

40. **Wikstrom, M., A. Jasaitis, B. C, A. Puustinen, and M. I. Verkhovsky**. 2000. The role of the D- and K-pathways of proton transfer in the function of the haem-copper oxidases. Biochim. Biophys. Acta **1459**:514-520.

41. **Ådelroth, P., M. Karpeforsa, G. Gildersona, F. L. Tomsonc, R. B. Gennis, and P. Brzezinski**. 2000. Proton transfer from glutamate 286 determines the transition rates between oxygen intermediates in cytochrome c oxidase. Biochim. Biophys. Acta **1459**:533-539.

42. **Branden, M., F. Tomson, R. B. Gennis, and P. Brzezinski**. 2002. The entry point of the K-proton-transfer pathway in cytochrome c oxidase. Biochemistry **41**:10794-10798.

43. **Tomson, F. L., J. E. Morgan, G. Gu, B. Barquera, T. V. Vygodina, and R. B. Gennis**. 2003. Substitutions for glutamate 101 in subunit II of cytochrome c oxidase from Rhodobacter sphaeroides result in blocking the proton-conducting K-channel. Biochemistry **42**:1711-1717.

44. **Junemann, S., B. Meunier, R. B. Gennis, and P. R. Rich**. 1997. Effects of mutation of the conserved lysine-362 in cytochrome c oxidase from Rhodobacter sphaeroides. Biochemistry **36**:14456-14464.

45. **Wikstrom, M., M. I. Verkhovsky, and G. Hummer**. 2003. Water-gated mechanism of proton translocation by cytochrome c oxidase. Biochim. Biophys. Acta **1604**:61-65.

46. **Zheng, X., D. M. Medvedev, J. Swanson, and A. A. Stuchebrukhov**. 2003. Computer simulation of water in cytochrome c oxidase. Biochim. Biophys. Acta **1557**:99-107.

47. **Olkhova, E., M. C. Hutter, M. A. Lill, V. Helms, and H. Michel**. 2004. Dynamic water networks in cytochrome c oxidase from *Paracoccus denitrificans* investigated by molecular dynamics simulations. Biophys. J. **86**:1873-1889.

.

48. **Ruitenberg, M., A. Kannt, E. Bamberg, B. Ludwig, H. Michel, and K. Fendler**. 2000. Single electron reducton of the oxidized state is coupled to proton uptake via the K pathway in *Paracoccus denitrificans* cytochrome c oxidase. Proc. Natl. Acad. Sci. USA **97**:4632-4636.

49. **Verkhovsky, M. I., A. Tuukkanen, C. Backgren, A. Puustinen, and M. Wikstrom**. 2001. Charge translocation coupled to electron injection into oxidized cytochrome c oxidase from *Paracoccus denitrificans*. Biochemistry **40**:7077-7083.

50. **Giuffre, A., M. C. Barone, M. Brunori, E. D'Itri, B. Ludwig, F. Malatesta, H. W. Muller, and P. Sarti**. 2002. Nitric oxide reacts with the single-electron reduced active site of cytochrome c oxidase. J. Biol. Chem. **277**:22402-22406.

51. **Forte, E., F. M. Scandurra, O. M. H. Richter, E. D'Itri, P. Sarti, M. Brunori, B. Ludwig, and A. Giuffre**. 2004. Proton uptake upon anaerobic reduction of the *Paracoccus denitrificans* cytochrome c oxidase: A kinetic investigation of the K354M and D124N mutants. Biochemistry **43**:2957-2963.

52. **Michel, H.** 1998. The mechanism of proton pumping by cytochrome c oxidase. Proc. Natl. Acad. Sci. USA **95**:12819-12824.

53. **Michel, H.** 1999. Cytochrome c oxidase: catalytic cycle and mechanisms of proton pumping-a discussion. Biochemistry **38**:15129-15140.

54. **Moody, A. J., and P. R. Rich**. 1990. The effect of pH on redox titrations of Heme a in cyanide liganded cytochrome c oxidase - Experimental and modeling studies. Biochim. Biophys. Acta **1015**:205-215.

55. Bockris, J. O. M., and A. K. N. Reddy (1973) *Modern Electrochemistry*. Vol. 1, Plenum, New York.

.

56. **Kassner, R. J.** 1973. A theoretical model for the effects of local nonpolar heme environments on the redox potentials in cytochromes. J. Am. Chem. Soc. **95**:2674-2676.

57. **Warshel, A., and S. T. Russell**. 1984. Calculations of electrostatic interactions in biological systems and in solutions. Q. Rev. Biophys. **17**:283-422.

58. **Rashin, A. A., and B. Honig**. 1985. Reevaluation of the Born model of ion hydration. J. Phys. Chem. **89**:5588-5593.

59. **Kim, J., J. Mao, and M. R. Gunner**. 2005. Are acidic and basic groups in buried proteins predicted to be ionized? J. Mol. Biol. **348**:1283-1298.

60. **Churg, A. K., and A. Warshel**. 1986. Control of the redox potential of cytochrome c and microscopic dielectric effects in proteins. Biochemistry **25**:1675-1681.

61. **Gunner, M. R., and B. Honig**. 1991. Electrostatic control of midpoint potentials in the cytochrome subunit of the *Rhodopseudomonas viridis* reaction center. Proc. Natl. Acad. Sci. USA **88**:9151-9155.

62. **Voigt, P., and E. W. Knapp**. 2003. Tuning heme redox potentials in the cytochrome c subunit of photosynthetic reaction centers. J. Biol. Chem. **278**:51993-52001.

63. **Mao, J., K. Hauser, and M. R. Gunner**. 2003. How cytochromes with different folds control heme redox potentials. Biochemistry **42**:9829-9840.

64. **Reedy, C. J., and B. R. Gibney**. 2004. Heme protein assemblies. Chem. Rev. **104**:617-649.

65. **Fann, Y. C., I. Ahmed, N. J. Blackburn, J. S. Boswell, M. L. Verkhovskaya, B. M. Hoffman, and M. Wikstrom**. 1995. Structure of Cu_B in the binuclear heme-copper center of the cytochrome aa_3-type quinol oxidase from *Bacillus subtilis*: an ENDOR and EXAFS study. Biochemistry **34**:10245-10255.

.

66. **Kannt, A., C. R. D. Lancaster, and H. Michel**. 1998. The coupling of electron transfer and proton translocation: electrostatic calculations on *Paracoccus denitrificans* cytochrome c oxidase. Biophys. J. **74**:708-721.

67. **Branden, M., A. Namslauer, O. Hansson, R. Aasa, and P. Brzezinski**. 2003. Water-hydroxide exchange reactions at the catalytic site of heme-copper oxidases. Biochemistry **42**:13178-13184.

68. **Puustinen, A., and M. Wikstrom**. 1999. Proton exit from the heme-copper oxidase of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **96**:35-37.

69. **Mills, D. A., and S. Ferguson-Miller**. 2002. Influence of structure, pH and membrane potential on proton movement in cytochrome oxidase. Biochim. Biophys. Acta **1555**:96-100.

70. **Behr, J., H. Michel, W. Mantele, and P. Hellwig**. 2000. Functional properties of the heme propionates in cytochrome c oxidase from *Paracoccus denitrificans*. Evidence from FTIR difference spectroscopy and site directed mutagenesis. J. Am. Chem. Soc. **39**:1356-1363.

71. **Wikstrom, M., A. Bogachev, M. Finel, J. E. Morgan, A. Puustinen, M. Raitio, M. Verkhovskaya, and M. I. Verkhovsky**. 1994. Mechanism of proton translocation by the respiratory oxidases. The histidine cycle. Biochim. Biophys. Acta **1187**:106-111.

72. **Popovic, D. M., and A. A. Stuchebrukhov**. 2004. Electrostatic study of the proton pumping mechanism in bovine heart cytochrome c oxidase. J. Am. Chem. Soc. **126**:1858-1871.

73. **Quenneville, J., D. M. Popovic, and A. A. Stuchebrukhov**. 2004. Redox-dependent pK_a of Cu_B histidine ligand in cytochrome c oxidase. J. Phys. Chem. B **108**:18383-18389.

74. **Mitchell, R., and P. R. Rich**. 1994. Proton uptake by cytochrome *c* oxidase on reduction and on ligand binding. Biochim. Biophys. Acta **1186**:19-26.

.

75. **Forte, E., M. C. Barone, M. Brunori, P. Sarti, and A. Giuffre**. 2002. Redox-linked protonation of cytochrome c oxidase: the effect of chloride bound to Cu_B. Biochemistry **41**:13046-13052.

76. **Hellwig, P., S. Grzybek, J. Behr, B. Ludwig, H. Michel, and W. Mantele**. 1999. Electrochemical and ultraviolet/visible/infrared spectroscopic analysis of heme a and a_3 redox reactions in the cytochrome c oxidase from *Paracoccus denitrificans*: separation of heme a and a_3 contributions and assignment of vibrational modes. Biochemistry **38**:1685-1694.

77. **Babcock, G. T., L. E. Vickery, and G. Palmer**. 1978. The electronic state of heme in cytochrome oxidase II. Oxidation-reduction potential interactions and heme iron spin state behavior observed in reductive titrations. J. Biol. Chem. **253**:2400-2411.

78. **Carithers, R. P., and G. Palmer**. 1981. Characterization of the potentiometric behavior of soluble cytochrome oxidase by magnetic circular dichroism. Evidence in support of heme-heme interaction. J. Biol. Chem. **256**:7967-7976.

79. Levinthal, C. (1969) How to fold graciously, in *Mossbauer Spectroscopy of Biological Systems* (Debrunner, P., Tsibris, J. C. M., and Münch, E., Eds.), U. Illinois Press, Urbana, Illinois.

80. **Zwanzig, R., A. Szabo, and B. Bagchi**. 1992. Levinthal's paradox. Proc Natl Acad Sci U S A **89**:20-22.

81. **Shurki, A., and A. Warshel**. 2003. Structure/function correlations of proteins using MM, QM/MM, and related approaches: methods, concepts, pitfalls, and current progress. Adv. Prot. CHem. **66**:249-313.

82. **Friesner, R. A., and V. Guallar**. 2005. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysi. Annu. Rev. Phys. Chem. **56**:389-427.

.

83. **Li, G., and Q. Cui**. 2003. pK_a calculations with QM/MM free energy perturbations. J. Phys. Chem. B **107**:14521-14528.

84. **Li, H., A. W. Hains, J. E. Everts, A. D. Robertson, and J. H. Jensen**. 2002. The prediction of protein pK_a's using QM/MM: the pK_a of lysine 55 in turkey ovomucoid third domain. J. Phys. Chem. B **106**:3486-3494.

85. **Riccardi, D., P. Schaefer, and Q. Cui**. 2005. pK_a calculations in solution and proteins with QM/MM free energy perturbation simulations: a quantitative test of QM/MM protocols. J. Phys. Chem. B **109**:17715 -17733.

86. **Jensen, J. H., H. Li, A. D. Robertson, and P. A. Molina**. 2005. Prediction and rationalization of protein pK_a values using QM and QM/MM methods. J. Phys. Chem. A. **109**:6634-6643.

87. **Schaefer, P., D. Riccardi, and Q. Cui**. 2005. Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules. J. Chem. Phys. **123**:14905.

88. **Honig, B., K. Sharp, and A.-S. Yang**. 1993. Macroscopic Models of Aqueous Solutions: Biological and chemical Applications. J. Phys. Chem. **97**:1101-1109.

89. **Schutz, C. N., and A. Warshel**. 2001. What are the "dielectric constants" of proteins and how to validate electrostatic models? Proteins: Struct. Funct. Genet. **44**:400-417.

90. **Ma, B. Y., M. Shatsky, H. J. Wolfson, and R. Nussinov**. 2002. Multiple diverse ligands binding at a single protein site: A matter of pre-existing populations. Protein Science **11**:184-197.

91. **Fradera, X., X. De La Cruz, C. H. Silva, J. L. Gelpi, F. J. Luque, and M. Orozco**. 2002. Ligand-induced changes in the binding sites of proteins. Bioinformatics **18**:939-948.

92. **Fields, P. A.** 2001. Review: Protein function at thermal extremes: balancing stability and flexibility. Comp Biochem Physiol A Mol Integr Physiol **129**:417-431.

.

93. **Perutz, M. F.** 1978. Electrostatic effects in proteins. Science **201**:1187-1191.

94. **Sharp, K., M. Gilson, R. Fine, and B. Honig**. 1987. Electrostatic interactions in proteins. Protein structure and Design **2**:235-244.

95. **Warshel, A., and J. Åqvist**. 1991. Electrostatic energy and macromolecular function. Ann. Rev. Biophys. Chem. **20**:267-298.

96. **Nakamura, H.** 1996. Roles of electrostatic interactions in proteins. Quart. Rev. Biophys. **29**:1-90.

97. **Matthew, J. B.** 1985. Electrostatic effects in proteins. Ann. Rev. Biophys. Chem. **14**:387-417.

98. **Feig, M., A. Onufriev, M. S. Lee, W. Im, D. A. Case, and C. L. Brooks III**. 2004. Performance comparison of Generalized Born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. J. Comput. Chem. **30**:265-284.

99. **Baker, N. A.** 2005. Improving implicit solvent simulations: a Poisson-centric view. Cur. Opin. Struct. Biol. **15**:137-143.

100. **Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. F. Shimanouchi, and M. Tasumi**. 1977. The protein data bank: A computer based archival file for macromolecular structures. J. Mol. Biol. **112**:535-542.

101. **Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne**. 2000. The protein data bank. Nucl. Ac. Res. **28**:235-242.

102. **Warshel, A.** 2002. Molecular dynamics simulations of biological reactions. Acc Chem Res **35**:385-395.

103. **Alexov, E. G., and M. R. Gunner**. 1997. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. Biophys. J. **72**:2075-2093.

.

104. **Mao, J., Y. Song, and M. R. Gunner**. In preparation. Better rotamer packing for continuum electrostatics pK_a calculations: MCCE2. J. Comp. Chem.

105. **Nicholls, A., and B. Honig**. 1991. A rapid finite difference algorithm utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. J. Comp. Chem. **12**:435-445.

106. **Richarz, R., and K. Wüthrich**. 1975. Carbon-13 NMR chemical shifts of the common amino acid residues measured in aqueous solutions of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. Biopolymers **17**:2133-2141.

107. **Matthew, J. B., F. R. N. Gurd, B. Garcia-Moreno, M. A. Flanagan, K. L. March, and S. J. Shire**. 1985. pH-dependent processes in proteins. CRC Crit. Rev. Biochem. **18**:91-197.

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

109. **Guex, N., and M. C. Peitsch**. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis **18**:2714-2723.

110. **Spassov, V. Z., H. Luecke, K. Gerwert, and D. Bashford**. 2001. pK_a alculations suggest storage of an excess proton in a hydrogen- bonded water network in bacteriorhodopsin. J Mol Biol **312**:203-219.

111. **Baasov, T., and M. Sheves**. 1986. Alteration of pKa of the Bacteriorhodopsin Protonated Schiff Base. A Study with Model Compounds. Biochemistry **25**:5249-5258.

112. **Georgescu, R. E., E. G. Alexov, and M. R. Gunner**. 2002. Combining conformational flexibility and continuum electrostatics for calculating pK_as in proteins. Biophys. J. **83**:1731-1748.

.

113. **Svensson-Ek, M., J. Abramson, G. Larsson, S. Tornroth, P. Brzezinski, and S. Iwata**. 2002. The X-ray crystal structures of wild-type and EQ(I-286) mutant cytochrome c oxidases from *Rhodobacter sphaeroides*. J. Mol. Biol. **321**:329-339.

114. **Song, Y., J. Mao, and M. R. Gunner**. 2003. Calculation of proton transfers in bacteriorhodopsin bR and M intermediates. Biochemistry **42**:9875-9888.

115. Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople. (1998) Gaussian 98, Revision A.9, Gaussian, Inc., Pittsburgh PA.

116. **Becke, A. D.** 1993. Density-Functional Thermochemistry .3. The role of exact exchange. J. Chem. Phys. **98**:5648-5652.

117. **Hay, P. J., and W. R. Wadt**. 1985. Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. J. Chem. Phys. **82**:270-283.

118. **Breneman, C. M., and K. B. Wiberg**. 1990. Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials - the Need for High Sampling Density in Formamide Conformational-Analysis. J. Comp. Chem. **11**:361-373.

.

119. **Song, Y., J. Mao, and M. R. Gunner**. in press. Electrostatic environment of hemes in proteins: pK_as of hydroxyl ligands. Biochemistry.

120. **Vanderkooi, G., and E. Stotz**. 1965. Reductive alteration of heme a hemochromes. J. Biol. Chem. **240**:3418-3424.

121. **Vanderkooi, G., and E. Stotz**. 1966. Oxidation-reduction potentials of heme a hemochromes. J. Biol. Chem. **241**:3316-3323.

122. **Battistuzzi, G., M. Borsari, J. A. Cowan, A. Ranieri, and M. Sola**. 2002. Control of cytochrome c redox potential: axial ligation and protein environment effects. J. Am. Chem. Soc. **124**:5315-5324.

123. **Wilson, G. S.** 1974. Electrochemical studies of porphyrin redox reactions as cytochromes models. Bioelectrochem. and Bioenerg. **1**:172-179.

124. **Marques, H. M., I. Cukrowski, and P. R. Vashi**. 2000. Co-ordination of weak field ligands by N-acetylmicroperoxidase-8 (NAcMP8), a ferric haempeptide from cytochrome c, and the influence of the axial lignad on the reduction potential of complexes of NAcMP8. J. Chem. Soc. **2000**:1335-1342.

125. **Santucci, R., H. Reinhard, and M. Brunori**. 1988. Direct electrochemistry of the undacapeptide from cytochrome c (microperoxidase) at a glassy carbon electrode. J. Am. Chem. Soc. **110**:8536-8357.

126. **Munro, O. Q., and H. M. Marques**. 1996. Heme-peptide models for hemoproteins. 1. solution chemistry of N-acetylmicroperoxidase-8. Inorg. Chem. **35**:3752-3767.

127. **Vashi, P. R., and H. M. Marques**. 2004. The coordination of imidazole and substituted pyridines by the hemeoctapeptide N-acetyl-ferromicroperoxidase-8 (Fe^{II}NAcMP8). J. Inorg. Biochem. **98**:1471-1482.

.

128. **Anderegg, G., and V. Gramlich**. 1994. Metal complexes of bivalent cobalt, nickel, copper, Zink, and cadmium with the tripodal ligand tris[2-(dimethylamino)ethyl]amine: their stabilities and the X-Ray crystal structure of its copper(T1) complex sulfate. Helv. Chim. Acta **77**:685-690.

129. **Thaler, F., C. D. Hubbard, F. W. Heinemann, R. v. Eldik, S. Schindler, I. Fabian, A. M. Dittler-Klingemann, F. E. Hahn, and C. Orvig**. 1998. Structural, spectroscopic, thermodynamic and kinetic properties of copper(II) complexes with tripodal tetraamines. Inorg. Chem. **37**:4022-4029.

130. **Sitter, A. J., J. R. Shifflett, and J. Terner**. 1988. Resonance Raman spectroscopic evidence for heme iron-hydroxide ligation in peroxidase alkaline forms. J. Biol. Chem. **263**:13032-13038.

131. **Peisach, J., W. E. Blumberg, S. Ogawa, E. A. Rachmilewitz, and R. Oltzik**. 1971. The effects of protein conformation on the heme symmetry in high spin ferric heme proteins as studied by electron paramagnetic resonance. J Biol Chem **246**:3342-3355.

132. **Song, Y., E. Michonova-Alexova, and M. R. Gunner**. in press. Calculated proton uptake on anaerobic reduction of cytochrome c oxidase: Is the reaction electroneutral? Biochemistry.

133. **Honig, B., and A. Nicholls**. 1995. Classical electrostatics in biology and chemistry. Science **268**:1144-1149.

134. **Gunner, M. R., and E. Alexov**. 2000. A pragmatic approach to structure based calculation of coupled proton and electron transfer in proteins. Biochim. Biophys. Acta **1458**:63-87.

135. **Fahmy, K., O. Weidlich, M. Engelhard, J. Tittor, D. Oesterhelt, and F. Siebert**. 1992. Identification of the Proton Acceptor of Schiff Base Deprotonation in Bacteriorhodopsin: A Fourier-Transform-Infrared Study of the Mutant Asp85 -> Glu in Its Natural Lipid Environment. Photochem. Photobiol. **56**:1073-1083.

.

136. **Lewis, A., J. Spoonhower, R. A. Bogomolni, R. H. Lozier, and W. Stoeckenius**. 1974. Tunable laser resonance raman spectroscopy of bacteriorhodopsin. Proc Natl Acad Sci U S A **71**:4462-4466.

137. **Zimanyi, L., G. Varo, M. Chang, B. Ni, R. Needleman, and J. K. Lanyi**. 1992. Pathways of proton release in the bacteriorhodopsin photocycle. Biochemistry **31**:8535-8543.

138. **Rammelsberg, R., G. Huhn, M. Lubben, and K. Gerwert**. 1998. Bacteriorhodopsin's intramolecular proton-release pathway consists of a hydrogen-bonded network. Biochemistry **37**:5001-5009.

139. **Gunner, M. R., M. A. Saleh, E. Cross, A. ud-Doula, and M. Wise**. 2000. Backbone dipoles generate positive potentials in all proteins: origins and implications of the effect. Biophys. J. **78**:1126-1144.

140. **Metz, G., F. Siebert, and M. Engelhard**. 1992. Asp85 is the only internal aspartic acid that gets protonated in the M intermediate and the purple-to-blue transition of bacteriorhodopsin. A solid-state 13C CP-MAS NMR investigation. FEBS Lett **303**:237-241.

141. **Fahmy, K., O. Weidlich, M. Engelhard, H. Sigrist, and F. Siebert**. 1993. Aspartic acid-212 of bacteriorhodopsin is ionized in the M and N photocycle intermediates: an FTIR study on specifically ^{13}C-labeled reconstituted purple membranes. Biochemistry **32**:5862-5869.

142. **Alexov, E. G., and M. R. Gunner**. 1999. Calculated protein and proton motions coupled to electron transfer: electron transfer from Q_A^- to Q_B in bacterial photosynthetic reaction centers. Biochemistry **38**:8253-8270.

143. **Subramniam, S., T. Marti, and H. G. Khorana**. 1990. Protonation state of ASP (Glu)-85 regulates the purple-to-blue transition in bacteriorhodopsin mutants Arg-82 --> Ala and Asp-85 --> Glu: The blue form is inactive in proton translocation. Proc. Natl. Acad. Sci. **87**:1013-1017.

.

144. **Balashov, S. P., R. Govindjee, E. S. Imasheva, S. Misra, T. G. Ebrey, Y. Feng, R. K. Crouch, and D. R. Menick**. 1995. The two pKa's of aspartate-85 and control of thermal isomerization and proton release in the arginine-82 to lysine mutant of bacteriorhodopsin. Biochemistry **34**:8820-8834.

145. **Balashov, S. P., E. S. Imasheva, R. Govindjee, and T. G. Ebrey**. 1996. Titration of aspartate-85 in bacteriorhodopsin: what it says about chromophore isomerization and proton release. Biophys. J. **70**:473-481.

146. **Druckmann, S., M. Ottolenghi, A. Pande, J. Pande, and R. H. Callender**. 1982. Acid-base equilibrium of the Schiff base in bacteriorhodopsin. Biochemistry **21**:4953-4959.

147. **Balashov, S. P., E. S. Imasheva, T. G. Ebrey, N. Chen, D. R. Menick, and R. K. Crouch**. 1997. Glutamate-194 to cysteine mutation inhibits fast light-induced proton release in bacteriorhodopsin. Biochemistry **36**:8671-8676.

148. **Dioumaev, A. K., H. T. Richter, L. S. Brown, M. Tanio, S. Tuzi, H. Saito, Y. Kimura, R. Needleman, and J. K. Lanyi**. 1998. Existence of a proton transfer chain in bacteriorhodopsin: participation of Glu-194 in the release of protons to the extracellular surface. Biochemistry **37**:2496-2506.

149. **Essen, L., R. Siegert, W. D. Lehmann, and D. Oesterhelt**. 1998. Lipid patches in membrane protein oligomers: crystal structure of the bacteriorhodopsin-lipid complex. Proc. Natl. Acad. Sci. USA **95**:11673-11678.

150. **Brown, L. S., J. Sasaki, H. Kandori, A. Maeda, R. Needleman, and J. K. Lanyi**. 1995. Glutamic acid 204 is the terminal proton release group at the extracellular surface of bacteriorhodopsin. J Biol Chem **270**:27122-27126.

.

151. **Kono, M., S. Misra, and T. G. Ebrey**. 1993. pH dependence of light-induced proton release by bacteriorhodopsin. FEBS Lett **331**:31-34.

152. **Onufriev, A., A. Smondyrev, and D. Bashford**. 2003. Proton affinity changes driving unidirectional proton transport in the bacteriorhodopsin photocycle. J. Mol. Biol. **332**:1183-1193.

153. **Luecke, H.** 2000. Atomic resolution structures of bacteriorhodopsin photocycle intermediates: the role of discrete water molecules in the function of this light-driven ion pump. Biochim Biophys Acta **1460**:133-156.

154. **Brown, L. S., and J. K. Lanyi**. 1996. Determination of the transiently lowered pKa of the retinal Schiff base during the photocycle of bacteriorhodopsin. Proc. Natl. Acad. Sci. USA **93**:1731-1734.

155. **Siegbahn, P. E. M., M. R. A. Blomberg, and M. L. Blomberg**. 2003. Theoretical study of the energetics of proton pumping and oxygen reduction in cytochrome oxidase. J. Phys. Chem. B **107**:10946-10955.

156. **Liao, G. L., and G. Palmer**. 1996. The reduced minus oxidized difference spectra of cytochromes a and a3. Biochim. Biophys. Acta **1274**:109-111.

157. **Siegbahn, P. E.** 2003. The catalytic cycle of tyrosinase: peroxide attack on the phenolate ring followed by O-O cleavage. J. Biol. Inorg. Chem. **8**:567-576.

158. **Proshlyakov, D. A., T. Ogura, K. Shinzawa-Itoh, S. Yoshikawa, and T. Kitagawa**. 1996. Resonance Raman/absorption characterization of the oxo intermediates of cytochrome c oxidase generated in its reaction with hydrogen peroxide: pH and H_2O_2 concentration dependence. Biochemistry **35**:8580-8586.

.

159. **MacMillan, F., A. Kannt, J. Behr, T. Prisner, and H. Michel**. 1999. Direct evidence for a tyrosine radical in the reaction of cytochrome c oxidase with hydrogen peroxide. Biochemistry **38**:9179-9184.

160. **Proshlyakov, D. A., M. A. Pressler, C. Demaso, J. F. Leykam, D. L. Dewitt, and G. T. Babcock**. 2000. Oxygen activation and reduction in respiration involvement of redox-active tyrosine 244. Science **290**:1588-1591.

161. **McCauley, K. M., J. M. Vrtis, J. Dupont, and W. A. van der Donk**. 2000. Insights into the functional role of the tyrosine-histidinde lindage in cytochrome c oxidase. J.Am.Chem. Soc.

162. **Boulatov, R., J. P. Collman, I. M. Shiryaeva, and C. J. Sunderland**. 2002. Functional analogues of the dioxygen reduction site in cytochrome oxidase: mechanistic aspects and possible effects of Cu_B. J. Am. Chem. Soc. **124**:11923-11935.

163. **Cappuccio, J. A., I. Ayala, G. I. Elliott, I. Szundi, J. Lewis, J. P. Konopelski, B. A. Barry, and O. Einarsdottir**. 2002. Modeling the active site of cytochrome oxidase: synthesis and characterization of a cross-linked histidine-phenol. J. Am. Chem. Soc. **124**:1750-1760.

164. **Pomes, R., G. Hummer, and M. Wikstrom**. 1999. Structure and dynamics of a proton shuttle in cytochrome c oxidase. Biochim. Biophys. Acta **1365**:255-260.

165. **Xu, J., and G. A. Voth**. 2005. Computer simulation of explicit proton translocation in cytochrome c oxidase: The D-pathway. Proc. Natl. Acad. Sci. USA **102**:6795–6800.

166. **Namslauer, A., A. Aagaard, A. Katsonouri, and P. Brzezinski**. 2003. Intramolecular proton-transfer reactions in a membrane-bound proton pump: the effect of pH on the peroxy to ferryl transition in cytochrome c oxidase. Biochemistry **42**:1488-1498.

.

167. **Wikstrom, M., C. Ribacka, M. Molin, L. Laakkonen, M. Verkhovsky, and A. Puustien**. 2005. Gating of proton and water transfer in the respiratory enzyme cytochrome c oxidase. Proc. Natl. Acad. Sci. USA **102**:10478-10481.

168. **You, T. J., and D. Bashford**. 1995. Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. Biophys. J. **69**:1721-1733.

169. **Bashford, D., and K. Gerwert**. 1992. Electrostatic calculations of the pK_a values of ionizable groups in bacteriorhodopsin. J. Mol. Biol. **224**:473-486.

170. **Bashford, D., and M. Karplus**. 1991. Multiple-site titration curves of proteins: An analysis of exact and approximate methods for their calculation. J. Phys. Chem. **95**:9556-9561.

171. **Antosiewicz, J., J. A. McCammon, and M. K. Gilson**. 1994. Prediction of pH-dependent properties of proteins. J.Mol.Biol. **238**:415-436.

172. **Antosiewicz, J., J. A. McCammon, and M. K. Gilson**. 1996. The determinants of pK_as in proteins. Biochemistry **35**:7819-7833.

173. **Honig, B. H., and W. L. Hubble**. 1984. Stability of "salt bridges" in membrane proteins. Proc. Natl. Acad. Sci. USA **81**:5412-5416.

174. **Mills, D. A., L. Florens, C. Hiser, J. Qian, and S. Ferguson-Miller**. 2000. Where is 'outside' in cytochrome c oxidase and how and when do protons get there? Biochim. Biophys. Acta **1458**:180-187.

175. **Branden, M., H. Sigurdson, A. Namslauer, R. B. Gennis, P. Adelroth, and P. Brzezinski**. 2001. On the role of the K-proton transfer pathway in cytochrome c oxidase. Proc. Natl. Acad. Sci. USA **98**:5013-5018.

.

176. **Verkhovsky, M. I., N. Belevich, J. E. Morgan, and M. Wikstrom**. 1999. Proton linkage of cytochrome a oxidoreduction in carbon monoxide-treated cytochrome c oxidase. Biochim. Biophys. Acta **1412**:184-189.

177. **Capitanio, N., G. Capitanio, M. Minuto, E. DeNitto, L. L. Palese, P. Nicholls, and S. Papa**. 2000. Coupling of electron transfer with proton transfer at heme a and Cu_A redox Bohr effects in cytochrome c oxidase. Biochemistry **39**:6373-6379.

178. **Kannt, A., U. Pfitzner, M. Ruitenberg, P. Hellwig, B. Ludwig, M. W., K. Fendler, and H. Michel**. 1999. Mutation of Arg-54 strongly influences heme composition and rate and directionality of electron transfer in *Paracoccus dinitrificans* cytochrome c oxidase. J.Biol.Chem. **274**:37974-37981.

179. **Mitchell, R., P. Mitchell, and P. R. Rich**. 1992. Protonation states of the catalytic intermediates of cytochrome c oxidase. Biochim. Biophys. Acta **1101**:188-191.

180. **Wikstrom, M., and M. I. Verkhovsky**. 2002. Proton translocation by cytochrome c oxidase in different phases of the catalytic cycle. Biochim Biophys Acta **1555**:128-132.

181. **Bloch, D., I. Belevich, A. Jasaitis, C. Ribacka, A. Puustinen, M. I. Verkhovsky, and M. Wikström**. 2004. The catalytic cycle of cytochrome c oxidase is not the sum of its two halves. Proc. Natl. Acad. Sci. USA **101**:529-533.

182. **Zhao, X., N. Yeung, Z. Wang, Z. Guo, and Y. Lu**. 2005. Effects of metal ions in the Cu_B on the redox properties of heme in heme-copper oxidases: spectroelectrochemical studies of an engineered heme-copper center in myoglobin. Biochemistry **44**:1210-1214.

183. **Seibold, S. A., D. A. Mills, S. Ferguson-Miller, and R. I. Cukier**. 2005. Water chain formation and possible proton pumping routes in *Rhodobacter sphaeroides* cytochrome c oxidase: A

.

molecular dynamics comparison of the wild type and R481K mutant. Biochemistry **44**:10475-10485.

184. **Olsson, M. H., P. K. Sharma, and A. Warshel**. 2005. Simulating redox coupled proton transfer in cytochrome c oxidase: looking for the proton bottleneck. FEBS Letts **579**:2026-2034.

185. **Hofacker, I., and K. Schulten**. 1998. Oxygen and proton pathways in cytochrome *c* oxidase. Proteins: Struct. Funct. Genet. **30**:100-107.

186. **Brunori, M., G. Amiconi, E. Antonin, J. Wyman, R. Zito, and A. R. Fanelli**. 1968. The transition between 'acid' and 'alkaline' ferric heme proteins. Biochim. Biophys. Acta **154**:315-322.

187. **Kraus, D. W., J. B. Wittenberg, J. F. Lu, and J. Peisach**. 1990. Hemoglobins of the *Lucina pectinata*/bacteria symbiosis. II. An electron paramagnetic resonance and optical spectral study of the ferric proteins. J. Biol. Chem. **265**:16054-16059.

188. **Sun, J., A. Wilks, P. R. Ortiz de Montellano, and T. M. Loehr**. 1993. Resonance Raman and EPR spectroscopic studies on heme-heme oxygenase complexes. Biochemistry **32**:14151-14157.

189. **Takahashi, S., J. Wang, D. L. Rousseau, K. Ishikawa, T. Yoshida, J. R. Host, and M. Ikeda-Saito**. 1994. Heme-heme oxygenase complex. Structure of the catalytic site and its implication for oxygen activation. J. Biol. Chem. **269**:1010-1014.

190. **Taylor, J. F.** 1981. Measurement of the oxidation-reduction equilibria of hemoglobin and myoglobin. Meth. Enz. **76**:577-582.

191. **Kraus, D. W., and J. B. Wittenberg**. 1990. Hemoglobins of the *Lucina pectinata*/bacteria symbiosis. I. Molecular properties, kinetics and equilibria of reactions with ligands. J. Biol. Chem. **265**:16043-16053.

.

192. **Yamada, H., R. Makino, and I. Yamazaki**. 1975. Effects of 2,4-substituents of deuteropheme upon redox potentials of horseradish peroxidases. Arch. Biochem. Biophys. **169**:344-353.

193. **Harbury, H. A.** 1957. Oxidation-reduction potentials of horseradish peroxidase. J. Biol. Chem. **225**:1009-1024.

194. **Sridharan, S., A. Nicholls, and B. Honig**. 1992. A new vertex algorithm to calculate solvent accessible surface areas. Biophys. J. **61**:A174.

195. **Weiss, M. S., A. Kreusch, E. Schiltz, U. Nestel, W. Welte, J. Weckesser, and G. E. Schulz**. 1991. The structure of porin from Rhodobacter capsulatus at 1.8 A resolution. FEBS Lett **280**:379-382.

196. **Jap, B. K., and P. J. Walian**. 1990. Biophysics of the structure and function of porins. Quarterly Review of Biophysics **23**:367-403.

197. **Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, J. Merz, K. M., D. M. Ferguson, D. C. Spellman, T. Fox, J. W. Caldwell, and P. A. Kollman**. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J. Am. Chem. Soc. **117**:5179-5197.

198. **Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen**. 1977. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. Journal of Computational Physics **23**:327-341.

199. **Beroza, P., D. R. Fredkin, M. Y. Okamura, and G. Feher**. 1991. Protonation of interacting residues in a protein by a Monte Carlo method: application to Lysozyme and the photosynthetic reaction center of *Rhodobacter sphaeroides*. Proc. Natl. Acad. Sci. USA **88**:5804-5808.

.