Cytochrome c_2 Exit Strategy: Dissociation Studies and Evolutionary Implications

Taras V. Pogorelov,[†] Felix Autenrieth,[†] Elijah Roberts,[‡] and Zaida A. Luthey-Schulten*[,†,‡]

Department of Chemistry and Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Chemical and Life Sciences Laboratory A544, MC-712, 600 South Mathews Avenue, Urbana, Illinois 61801

Received: August 2, 2006; In Final Form: October 17, 2006

Small, water-soluble, type c cytochromes form a transient network connecting major bioenergetic membrane protein complexes in both photosynthesis and respiration. In the photosynthesis cycle of *Rhodobacter sphaeroides*, cytochrome c_2 (cyt c_2) docks to the reaction center (RC), undergoes electron transfer, and exits for the cytochrome bc_1 complex. Translations of cyt c_2 about the RC−cyt c_2 docking interface and surrounding membrane reveal possible exit pathways. A pathway at a minimal elevation allowed by the architecture of the RC is analyzed using both an all-atom steered molecular dynamics simulation of the RC−cyt c_2 complex and a bioinformatic analysis of the structures and sequences of cyt c. The structure-based phylogenetic analysis allows for the identification of structural elements that have evolved to satisfy the requirements of having multiple functional partners. The patterns of evolutionary variation obtained from the phylogenetic analysis of both docking partners of cyt c_2 reveal conservation of key residues involved in the interaction interfaces that would be candidates for further experimental studies. Additionally, using the molecular mechanics Poisson−Boltzmann surface area method we calculate that the binding free energy of reduced cyt c_2 to the RC is nearly 6 kcal/mol more favorable than with oxidized cyt c_2. The redox-dependent variations lead to changes in structural flexibility, behavior of the interfacial water molecules, and eventually changes in the binding free energy of the complex.

1. Introduction

Type c cytochromes (c-cyt), small water-soluble redox proteins, participate in a network of transient complexes in the fundamental biological processes of photosynthesis and respiration.[1−4] In cyclic photosynthesis, the main energy source of purple bacteria, cyt c_2 shuttles electrons between two major integral membrane protein complexes of ubiquinone/cytochrome c oxidoreductase (cytochrome bc_1 complex (cyt bc_1)) and the photosynthetic reaction center (RC). In the respiratory pathway present in the mitochondria of all eukaryotic organisms as well as in many bacteria and archaea, cyt c shuttles electrons between cyt bc_1 and cytochrome c oxidase (cyt c oxidase). In all of these processes, the electron-transfer events are coupled to ATP synthesis.[5]

Since the publication of the first structure of a photosynthetic reaction center for *Rhodopseudomonas viridis*,[6] several high-resolution X-ray structures of electron-transfer complexes are now available: a detailed structure of cyt c_2 docked to the RC of *Rhodobacter* (*Rb.*) *sphaeroides*[2,7] (shown in Figure 1 embedded in a 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) membrane) and structures of cyt c docked to the cytochrome bc_1 complex[8,9] and cytochrome c peroxidase,[10,11] both from yeast *Saccharomyces cerevisiae*. Models of horse cyt c complexed to cyt c oxidase from *Paracoccus denitrificans*[15,16] and to bovine cyt c oxidase[12−14] also have been developed. The functional constraint of binding to multiple partners resulted in cyt c acquiring specific sequence and structural features. This evolutionary pressure has led to the existence of three classes of water-soluble cyt c, small, medium, and large, each with well-conserved sequences and insertions.[17,18] The SCOP database[38] contains over 40 cytochrome c structures (for a recent review, see ref 34).

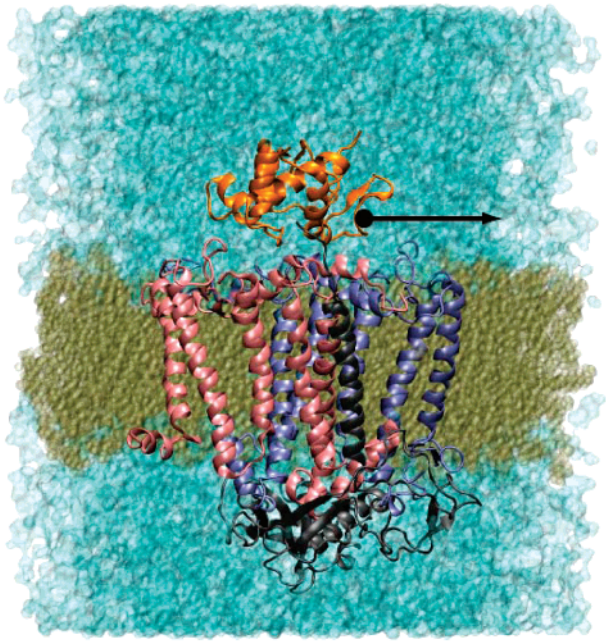

Figure 1. Simulated RC−cyt c_2 complex of *Rb. sphaeroides*. The L (blue), M (pink), and H (gray) RC subunits and cyt c_2 (orange) are embedded in a lipid bilayer (gold) and are shown with explicit water molecules (cyan). Oxidized cyt c_2 in the shown equilibrated complex was elevated to 3 Å and then pulled for 35 Å along the RC surface. The black arrow indicates the direction of cyt c_2 pulling away from its initial RC docking position.

* Author to whom correspondence should be addressed. E-mail: schulten@scs.uiuc.edu.
[†] Department of Chemistry.
[‡] Center for Biophysics and Computational Biology.

10.1021/jp064973i CCC: $37.00
Published on Web 12/23/2006

A primary characteristic of c-cyt docking interfaces in photosynthetic and respiratory complexes is the extensive presence of weak protein−protein interactions between cyt c and its respective docking partner. The transient nature of the redox complexes formed between c-cyt and its partners has led to a smaller interfacial contact area and a relatively low packing density of the interfacial residues.[19] The geometry of the contact domain is generally described as a small patch of hydrophobic residues surrounded by a domain of partially solvated charged residues. In some cases a cation−π pair is also present at the periphery. These features support the transient nature of cyt c docking and allow rapid electron-transfer and turnover rates in the docked complex.[2]

In a crystallographic study Alexrod et al.[7] described the binding interface of the RC−cyt c_2 complex in *Rb. sphaeroides* as being spatially divided into two distinct domains. On the RC one domain consists of hydrophobic residues involved in direct contact, a distance of less than 4 Å, to cyt c_2. This includes TyrL162 (here TyrL162 denotes the tyrosine residue 162 on the L subunit of the RC), which makes contact with the heme in cyt c_2 and several nearby interacting nonpolar groups including a cation−π pair. In the cation−π pair between ArgC32 (Arg32 of cyt c_2) and TyrM295 (Tyr295 of the M subunit of the RC) the plane of the aromatic ring of the tyrosine is parallel to and separated by less than 4 Å from the arginine side chain. The docked cyt c_2 straddles the boundary between the M and the L subunits of the RC. The long-range interaction domain, shaped as a half-donut mostly over the M subunit, surrounds the so-called short-range hydrophobic domain and consists of complementary charged residues, with cyt c_2 having the positive charges and the RC the negative charges. The side chains of most of the charged groups in the crystal structure are separated by more than 5 Å. The only notable exception is the AspL261−LysC99 pair, which is separated by 4.8 Å. Crystallographers have reported a number of water molecules, which they divide into two classes: the bridging water molecules, which form hydrogen bonds with both proteins, and water molecules that form hydrogen bonds with only one of the proteins. The fact that the charged residues at the interface are solvent-accessible and all oppositely charged residues are separated by more than 3.5 Å implies that no permanent salt bridges are formed.

In our previous molecular dynamics study of the undocking of cyt c_2 from the neutral RC of *Rb. sphaeroides*, we examined the behavior of water and residues at the interface.[20] The dynamics of the interfacial waters were dependent on the oxidation state of cyt c_2, and the formation of transient salt bridges between AspL261 and LysC97/LysC99 with a period of approximately 0.5 ns was reported. In the equilibration simulations of both docked redox states, most of the water molecules of the interface were exchanged with the bulk, including many of the water molecules reported in the X-ray structure. Interestingly, a redox-dependent behavior was observed for water molecules near the cation−π pair, ArgC32−TyrM295. During the length of the study for the complex with reduced cyt c_2 (cyt c_2^{red}, where the heme iron has a formal oxidation state of +2), at least four single-file water molecules remained in the hydrophobic pocket between the cation−π pair and residues involved in the electron-transfer (ET) pathway, TyrL162 and LeuM191. The high residence time of these water molecules suggests that they could participate in ET. In the complex with oxidized cyt c_2 (cyt c_2^{ox}, where the heme iron has a formal oxidation state of +3), the interfacial water molecules showed a much larger mobility. The calculations also showed that the position of the energy minimum shifts slightly upon oxidation of cyt c_2 allowing the protein to wobble and interfacial waters to more readily exchange. This greater mobility of the interfacial waters could potentially initiate undocking.

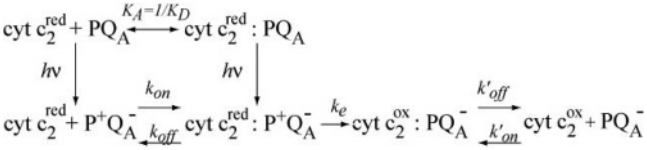

Figure 2. Reaction scheme used to interpret the observed biphasic kinetics.

While the binding constant of the complex with cyt c_2^{red} has been measured, there are still questions about the binding of cyt c_2^{ox}. Does it leave the same way cyt c_2^{red} came or are there other probable exit pathways? What residues are important for the undocking? What is the effect of the cytochrome's oxidation state on the binding constant of RC−cyt c_2?

Using both steered molecular dynamics (SMD) simulations and insights gained from our bioinformatics study, we argue here that an energetically favorable exit pathway exists along the L subunit and surrounding membrane. We also use the molecular mechanics Poisson−Boltzmann surface area (MM-PBSA) method to calculate binding free energies of the RC−cyt c_2 complex for both redox states of cyt c_2. An evolutionary analysis of cyt c_2 suggests that conserved residues and structural elements have coevolved to satisfy the functional constraint of binding to both the RC and the cytochrome bc_1 complex. This evolutionary analysis is used to interpret the results of our MD simulations and the plethora of mutagenesis data and to make a prediction about the binding affinity of RC−cyt c_2 in *Rhodospirillum* (*R.*) *rubrum*.

To motivate the bioinformatic study, we begin by reviewing the results of the mutagenesis experiments, which have shown substantial changes in the kinetics. These mutations are of highly conserved residues as seen in the alignments presented in the Results section. Mutations invoking little change are often observed to be present in related species. The key mutations are reviewed below, and the corresponding kinetic data are provided in the Supporting Information.

1.1. Mutagenesis Experiments. Effects of mutations on the electron-transfer rates and dissociation constants have been determined in a series of experiments in the RC−cyt c_2 complex of *Rb. sphaeroides* using transient absorption spectroscopy with flash photolysis.[21−25] The general two-state reaction scheme used to describe the binding and electron-transfer events is summarized in Figure 2.[21,26] Prior to laser photo-oxidation of the primary donor of the RC, cyt c_2^{red} and neutral RC exist in equilibrium between a free state and a bound state described by the dissociation constant $K_D = [PQ_A][\text{cyt } c_2^{red}]/[PQ_A-\text{cyt } c_2^{red}] \approx 10^{-6}$ M, under the condition of low ionic strength. Here PQ_A represents the neutral RC, where P is the primary donor, a special pair of bacteriochlorophylls, and Q_A is the bound ubiquinone at the end of the A electron-transfer branch, located in the L subunit of the RC. Upon laser flash charge separation occurs, and the primary donor P is oxidized to P^+. Decay generally exhibits biphasic kinetics: the fast first phase due to the electron transfer from already bound cyt c_2 and a slower second phase due to binding of cyt c_2 followed by the electron-transfer event. The second-order electron-transfer rate for the free cyt c_2 is $k_2 = k_e k_{on}/(k_{off} + k_e)$.[23]

Most of the mutagenesis measurements are conducted in the diffusion-limited regime, in which the dissociation rate $k_{off} = 10^3\ s^{-1}$, see ref 27, is much slower than the first-order electron-transfer rate $k_e = 10^6\ s^{-1}$, and thus the observed second-order

rate constant is approximately equal to the association constant, $k_2 \approx k_{on} \approx 10^9\ s^{-1}\ M^{-1}$, observed in the wild-type complex. In the diffusion-limited regime, the binding and reactivity are essentially decoupled, and the overall rate is independent of k_e. Some mutations dramatically alter the dissociation rate, such that $k_{off} \gg k_e$ was observed.[22] In this fast exchange reaction regime $k_2 \approx k_e/K_D$. Most of the recent experimental data have been interpreted using the reaction scheme just described.

1.1.1. Charged residues. An extensive study of the role of electrostatic interactions in the association of cyt c_2 with the RC of *Rb. sphaeroides* and electron transfer was performed using charge-modifying mutations.[23] The largest reduction in the binding affinity as measured by the increase in the dissociation constant K_D was observed for RC mutants with reduced negative charge. The largest decrease was reported for the AspM184Lys mutant and the double mutant AspL155Lys/AspL257Lys on the L subunit (Figure 11). Most of the negatively charged residues are asymmetrically located on the M subunit, which is the proposed location of the transition complex.[28] The authors also observed that the changes in the binding constants are correlated with the rate of collisional electron transfer, k_2. Finally, the free energy change for the mutation of a negative residue to a neutral residue was measured to be approximately one-half of the free energy change for the mutation of a negative residue to a positive residue. This suggests that the changes in the desolvation energy, required to remove solvent, and hydrophobic energies are small compared to changes in electrostatic energy. The authors interpreted these kinetic and thermodynamic studies as suggestive that charged residues are solvated and do not form salt bridges. An interesting exception was the mutation of AspL261 where the ratio of the free energy changes is 1:3, suggesting that the desolvation and/or hydrophobic contributions accompanying the reversal of charge are much larger. AspL261 was observed in our previous molecular dynamics study[20] to form a transient salt bridge with LysC97, which could explain the energetic differences.

1.1.2. Nonpolar residues. On the RC the three hydrophobic residues TyrL162, LeuM191, and ValM192 are in direct contact with cyt c_2; see Figure 11 for their location. Mutations of LeuM191 and TyrL162 to isoleucine and phenylalanine, respectively, are observed in many related species and have little effect on the kinetics as shown in the Supporting Information. While many of the mutations appear to leave the system still in the diffusion-limited regime, the TyrL162Ala and the LeuM191Lys mutations produced a 100-fold and 1000-fold increase in the dissociation constant, respectively.[22] Assuming little or no change in the association rate k_{on}, the corresponding decreases in k_e and k_2 would be consistent with an increase the dissociation rate k_{off}.

1.1.3. Cation−π Pair. The functional role of the cation−π interaction ArgC32−TyrM295 in the binding and electron transfer was recently addressed by both mutational studies[21] and MD simulations of the undocking process.[20] In the experimental studies, all of the single mutations of TyrM295 resulted in only a slight increase of K_D, whereas the replacement of the charged ArgC32 by alanine or the double mutant ArgC32Ala/TyrM265Ala increased K_D 80-fold and 100-fold, respectively. All mutants had a k_e close to the wild-type value, indicating that once cyt c_2 is bound its position is not significantly affected by the C32/M265 pair. The near-constant second-order rate of electron transfer strongly suggests that the cation−π interaction plays a negligible role in the rate of association k_{on} and that its primary role is in the undocking process described by k_{off}. The MD simulations showed that the breaking of the ArgC32−TyrM295 interaction is comparable energetically to the breaking of any of the transient salt bridges formed during the equilibration.[20] As we show in the Results section, the cation−π pair in *Rb. sphaeroides* is one of the last "bonds" to break as cyt c_2 is displaced along the exit pathway from the original docking site.

1.1.4. Redox-Dependent Variations. Plasmon-waveguide resonance (PWR) spectroscopy was used by Cusanovich and coworkers to directly measure the binding of the oxidized and reduced cyt c_2 of the closely related *Rhodobacter* (*Rb.*) *capsulatus* to the neutral RC from *Rb. sphaeroides* in a lipid bilayer.[29] The measurements were performed at both low (19 mM) and high or physiological (109 mM) ionic strengths on the wild-type and a LysC93Pro mutant of cyt c_2, which corresponds to LysC97Pro in *Rb. sphaeroides*. This residue is located at the beginning of a loop forming one of the major interface elements (Figure 9) and was observed to make transient salt bridges in the MD equilibration simulations. Most notably, the authors report that under low ionic strength conditions the dissociation constant for cyt c_2^{ox} is 30 times greater than that for cyt c_2^{red}. In addition they detected the existence of two binding sites of high and low binding affinity (K_D = 10 and 150 nM) on the RC. The high-affinity site is accessible only by wild-type cyt c_2^{red}, indicating that a shift in the binding site accompanies the oxidation of the cytochrome. Furthermore the increase in K_D for the mutated system under physiological conditions shows that the modification in the so-called hinge region, residues 88−102 (92−106 in *Rb. sphaeroides*), seems to reduce the stability of the complex. It was concluded that the dynamics of the hinge region are instrumental in the dissociation of cyt c_2^{ox}, as they alter the dissociation rate k_{off}. Details of the experimental results described above are summarized in Tables 1−4 of the Supporting Information.

1.2. Entrance Pathway. Recently it has been suggested that cyt c_2 reaches the active bound state by approaching the RC from the M subunit side, first forming an encounter complex located 10 Å away from the bound position.[30] In their study, Onuchic and co-workers performed a computational analysis of the kinetic study performed by Okamura and colleagues[31] to determine the ensemble of configurations for the encounter complex and transition state. The resulting transition state complex structures revealed strong electrostatic interactions between charged residues mostly located on the M subunit of the RC and cyt c_2, in particular between LysC99 and AspM184/GluM95. The only residues on the L subunit that were found to be important in the formation of the transition state complex were AspL257, AspL261, and AspL155 near the boundary with the M subunit. Additionally it was revealed that the orientation of cyt c_2 plays a role in the formation of the short-range van der Waals contacts that stabilize the active electron-transfer complex, which was also seen in the MD simulations.[20]

1.3. Evolution. The sequences of cyt c have been studied phylogenetically for more than 40 years.[18,32−34] Cytochrome c is an excellent metric for evolutionary change in Eukarya,[32] due to its presence in the mitochondrial respiratory chains, and for evolution within the subdivisions of Proteobacteria.[35] By comparative analysis Woese showed that phylogenetic trees derived from the analysis of the 16S (small subunit) ribosomal RNA sequences and cyt c sequences yield essentially the same evolutionary relationships among the purple bacteria.[18,35]

As protein structure is more highly conserved than sequence,[36] better alignments can be constructed by combining both sequence and structural information.[37] Over the past decade many cyt c structures have become available,[38] allowing us to revisit the cyt c phylogeny and correlate the changes in sequence

with specific structural elements. In this study we used a combination of sequence[39] and structural[37] bioinformatics methods to create more accurate alignments of the sequences of cyt c_2, the L and M subunits of the RC, and cyt c_1 of the cytochrome bc_1 complex. As the resulting sequence and structure phylogenetic trees for cyt c_2 are congruent, we can infer interactions in related organisms. Bioinformatics studies provide an opportunity to analyze interactions of cyt c_2 with the RC and with the cytochrome bc_1 complex. In the Results section of this paper we identify essential structure and sequence elements for the docking and undocking processes that complement the crystallographic and mutagenesis studies. Additionally, we use this information to predict the changes in the binding activity of cyt c_2 in *R. rubrum*. Finally, we construct and analyze a model of the cyt bc_1−cyt c_2 complex of *Rb. capsulatus*.

1.4. Dissociation of Cytochrome c_2. The binding of cyt c_2^{red} with the RC has been studied experimentally and computationally with various levels of detail. On the basis of the analysis of the electrostatic interactions, it has been suggested that association of cyt c_2^{red} is initiated from the M subunit. Here we look at the other very important part of the cycle, dissociation of cyt c_2^{ox} from the RC. We propose a dissociation pathway for cyt c_2^{ox} along the L subunit of the RC that is mostly devoid of complementary charged residues. We explore the proposed dissociation path in detail using a large-scale full-atom SMD simulation and compare the results to the nonbonded energies required to pull cyt c_2^{ox} across the surrounding membrane. The force and energy profiles obtained from the SMD simulations of the dissociation are discussed in the broader context of evolutionary analysis. Our bioinformatic analysis takes into account structural and sequence information from Bacteria and Eukarya as well as biochemical data from selected purple bacterial systems and helps us to understand how cyt c_2 structure accommodates docking to both the RC and the cytochrome bc_1 complex. Finally, using the molecular dynamics simulation data, we carry out MM-PBSA calculations of the binding free energy for reduced and oxidized cyt c_2 to a neutral RC.

2. Methods

2.1. System Setup. The RC−cyt c_2 complex for *Rb. sphaeroides* shown in Figure 1 was set up in a POPC membrane of hexagonal geometry with the TIP3 water model[40] as described in ref 20. Equilibration of the complex in both redox states was performed in the NPT ensemble with full particle mesh Ewald (PME) electrostatics and a cutoff of 12 Å for the van der Waals interactions. Periodic boundary conditions were applied with a flexible cell. The entire system of the RC, cyt c_2, water molecules, and ions is nearly 145 000 atoms in total. An integration step of 1 fs was used. Langevin dynamics was used to maintain a constant temperature at 298 K, and the hybrid Nosè−Hoover Langevin piston was used to maintain a constant pressure of 1 atm. The complex has the RC in a neutral state and cyt c_2 in an either oxidized or reduced state with the heme parameters as specified in ref 41 and CHARMM27.[42] A neutral RC with neutral special pairs is a model for the experimentally studied complexes.[7,23] The last nanosecond of the 4 ns equilibration was used for calculating the free energy of binding. The SMD simulation along the exit pathway was initiated with the final structure of the complex with the oxidized cyt c_2 at the end of the 4 ns simulation. All simulations were performed with the NAMD2 program[43,44] using the CHARMM27 force field.[42]

A simplified RC−cyt c_2 model of *R. rubrum*[45] for the electrostatic studies was generated in which all charged cytoplasmic parts of the RC involved in docking were used and the artificially created termini were rendered neutral. The *R. rubrum* cyt c_2 (PDB ID 3C2C) was docked to the RC model by superimposing its coordinates upon the RC−cyt c_2 X-ray structure (PDB ID 1L9B). The resulting RC−cyt c_2 model of *R. rubrum* was solvated using the SOLVATE feature of VMD,[46,47] and all water molecules further than 7 Å from the protein surface were removed. The resulting 20 000 atom system was minimized for 4000 steps and equilibrated for 0.2 ns in the NVT ensemble.

2.2. Energy Landscape of the Exit Pathway. *2.2.1. Initial Translational Scan.* To investigate the geometric fit between cyt c_2 and the RC along different pathways, cyt c_2 was translated along the RC surface at various elevations in the absence of water. Cyt c_2 was moved 20 Å in 0.5 Å steps along the horizontal plane in each direction from the docked position at elevations of 1, 2, and 3 Å above the RC. At each point the structure of the complex was minimized for 20 000 steps in NAMD2,[43,44] and the variation in van der Waals (VdW) energy of the RC−cyt c_2 complex with a cutoff of 12 Å was used to locate a probable exit pathway. The origin was set at the position derived from the minimized and equilibrated oxidized RC−cyt c_2 complex and is close to the position of the reduced complex defined in the X-ray structure (Figure 3).[7]

2.2.2. SMD Simulations along the Suggested Exit Pathway. Cytochrome c_2 was pulled along the detected exit pathway toward the L subunit for 35 Å. Forces were only applied to all backbone atoms of cyt c_2; the side chains and the heme group were allowed to move freely. The external force F was of the form

$$F = k(vt - \Delta x)$$

In this form Δx is the displacement of the center of mass of the pulled atoms relative to its original position, v is the constant velocity of a spring attached to the center of mass of the pulled atoms,[20,49] and k is the spring constant. The SMD simulation was performed, using NAMD2[43,44] in the NVT ensemble with a spring constant k of 500 pN/Å (7.196 kcal/(mol/Å^2)) and a pulling speed of 5 Å/ns for 7 ns. The time evolution of the force and the displacement of the cyt c_2 center of mass were recorded every 0.1 ps. The model system with cyt c_2 elevated to 3 Å over the membrane without the RC was prepared and modeled in a similar way. Visualization and analysis of the simulations were preformed using the program VMD.[46]

Van der Waals and electrostatic contributions between the RC and cyt c_2 over the SMD trajectories along the exit pathway were monitored using the program MDenergy in the NAMD2 package.[43,44] The electrostatic energies originally obtained by setting the dielectric constant to 1 were later scaled by a factor of 10 (see the Results section for more details) to approximate the screening effects of interface water. No cut-off factors were employed for either energy term.

Simulations were performed on the Pittsburgh Supercomputing Center's Lemieux terascale platform using 256 processors, the National Center for Supercomputing Application's (NCSA) Xeon Cluster (Tungsten) using 256 processors, and a local 127 processor SGI Origin 2000 cluster.

2.3. Electrostatic Surfaces. Visualization of the electrostatic potential isosurfaces for cyt c_2 and the RC was performed using the PME plugin[55] of VMD.[46] The electrostatic potential is obtained by numerically solving the Poisson equation for the system of charges approximated by spherical Gaussians. If the system under consideration is charged, then the PME plugin distributes the total charge over the grid points used in the calculation, which usually leads to very good approximations

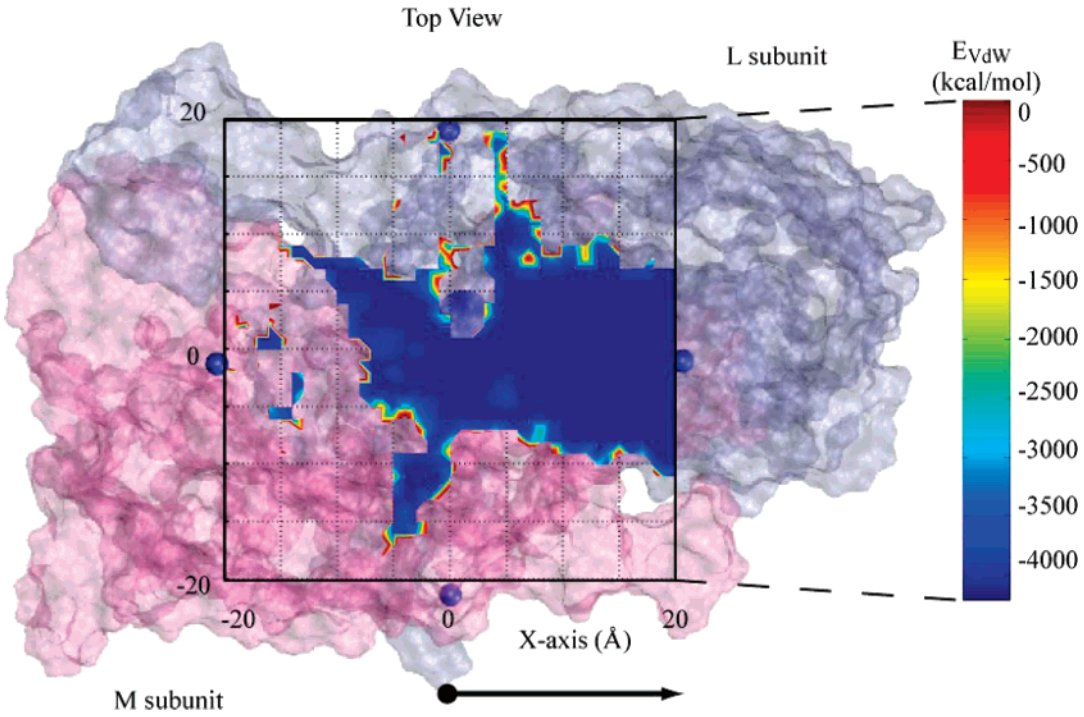

Figure 3. Surface of the RC overlaid with a grid showing the VdW interaction energies between cyt c_2 and the RC at an elevation of 3 Å. The most favorable energy values are depicted in blue. A possible horizontal dissociation pathway for cyt c_2 is revealed along the L subunit (translucent blue). Paths over the M subunit (translucent pink) do not appear to be favorable.

for weakly charged systems. The potential was calculated without ions, water molecules, or the membrane, and values were scaled according to DelPhi calculations on the same system. We performed comparative analysis for complexes of *Rb. sphaeroides* and *R. rubrum*.

2.4. Binding Free Energy: Continuum Electrostatic Model. The electrostatic contribution to the free energy of binding, $\Delta G_{elect}(RC-cyt, r_{sep})$, was calculated at different elevations (r_{sep}) of cyt c_2 from the RC

$$\Delta G_{elect}(RC-cyt, r_{sep}) = G_{elect}(RC-cyt, r_{sep}) - G_{elect}(RC, r_{sep}) - G_{elect}(cyt, r_{sep}) \quad (1)$$

where $G_{elect}(RC-cyt, r_{sep})$, $G_{elect}(RC, r_{sep})$, and $G_{elect}(cyt, r_{sep})$ are the electrostatic energies of the conformations of the complex and individual subunits obtained at each vertical separation from the SMD separation.

According to the continuum electrostatic model, the electrostatic free energies of each system arise from two types of interactions

$$\Delta G_{elect} = \Delta E_{Coul}(\epsilon_{in}) + \Delta G_{solv}^{polar}(\epsilon_{in}, \epsilon_{out}) \quad (2)$$

E_{Coul} is the pairwise Coulombic energy calculated between charges on each protein using the value of the internal dielectric constant, ϵ_{in}. The polar solvation energy, $\Delta G_{solv}^{polar}(\epsilon_{in}, \epsilon_{out})$, in the absence of salt, is the energy of transferring the system from an environment with a dielectric constant equal to ϵ_{in} (typically 4 for proteins) to the one with a dielectric constant characteristic of the solvent, ϵ_{out}, here 80 for water. Both contributions were obtained by numerically solving the Poisson−Boltzmann equation (PBE) using the program DelPhi, version 4,[50,51] with two grid points per ångström and the box size set so that the system filled 80% of the box. No membrane lipids were included in the calculations. Inclusion of a small number of water molecules with high residency time caused only small changes in the free energies (data not shown); thus no water molecules were included. Parameters for solvation energy (PARSE) charges and radii were used in the DelPhi calculation[53,54] along with $\epsilon_{in} = 4$. All calculations were performed over the equilibration runs.

2.5. Binding Free Energy: Molecular Mechanics Poisson−Boltzmann Surface Area Method. The binding free energies of the complexes are estimated using the MM-PBSA method developed by Kollman and co-workers[56,57] and recently extended to calculate the dimerization free energy of insulin dimers.[58] In this method, one uses structures obtained from MD simulations of the complexes to calculate the energies of the interaction at each step in the trajectory. These values are averaged over a simulation to give a free energy of binding. The free energy of interaction is the difference of the free energy of the complex and its components, excluding water, ions, and the membrane

$$\Delta G_{binding} = G^{complex} - G^{RC} - G^{cyt\ c_2} \quad (3)$$

where

$$G = E_{MM} + G_{solv} - TS \quad (4)$$

Here the molecular mechanical energy is

$$E_{MM} = E_{geom} + E_{Coul} + E_{VdW} \quad (5)$$

where E_{geom} includes energy terms that define the geometry of the molecule, i.e., bonds, angles, and dihedral angles, E_{Coul} is the pairwise Coulomb energy between the charges assuming $\epsilon = 1$, and E_{VdW} is the van der Waals interaction energy. The nonbonded energies were calculated from MD trajectories using the MDenergy program with CHARMM27 parameters.

The solvation free energy has both polar and nonpolar contributions

$$G_{solv} = G_{solv}^{polar}(\epsilon_{in}, \epsilon_{out}) + G_{solv}^{nonpolar} \quad (6)$$

As in the continuum electrostatic model, the polar contribution was calculated by solving the Poisson−Boltzmann equation using the program DelPhi,[50,51] except with the charge parameters and atomic radii from CHARMM27.[41,52] The interior dielectric constant was set to 1, and the dielectric constant of the exterior to 80.

Nonpolar contribution to the solvation free energy is due to the formation of a cavity for the solute and is estimated from the solvent-accessible surface area (SASA) of the proteins using the empirical formula[53]

$$G_{\mathrm{solv}}^{\mathrm{nonpolar}} = \gamma \mathrm{SASA} + b \quad (7)$$

where $\gamma = 0.00542$ kcal mol^{-1} Å^{-2} and $b = 0.92$ kcal mol^{-1}. The SASA, which in the crystal structure is 1280 Å^2, was calculated using VMD at each step.

Neglecting the contributions from the solvent, the entropic contributions are taken to be the sum of translational, S_{trans}, rotational, S_{rot}, and vibrational, S_{vib}, degrees of freedom of the proteins

$$S = S_{\mathrm{trans}} + S_{\mathrm{rot}} + S_{\mathrm{vib}} \quad (8)$$

As there was little change in geometry between the reduced and the oxidized cyt c_2, the contributions were assumed to be the same for both systems. Formation of the complex leads to a decrease in the rotational and translational degrees of freedom. While there is a general agreement about the trends, the magnitude of the entropic contributions is still debated in the literature,[57,58,60−63] and we chose to follow the approach first introduced by Finkelstein and Janin in which the entropy contributions of translational and rotational degrees of freedom for cyt c_2 are estimated from the statistical thermodynamics of ideal gas and solution,[63,64] e.g., $S_{\mathrm{trans}}^{\mathrm{ideal}} = R \ln[\pi^{3/2}\, \mathrm{e}^{3/2}(V/N)\lambda^3]$, where $\lambda = h/\sqrt{2mk_{\mathrm{B}}T}$ is the thermal wavelength of a particle of mass m, $k_{\mathrm{B}}T$ is thermal energy, N is the number of particles, and V is the volume of the solution. The change in the translational entropy can be estimated from the motions of the molecules in the crystal structures

$$\Delta S_{\mathrm{trans}} \approx R \ln(\delta x_1 \delta x_2 \delta x_3 / V_0) \approx n_{\mathrm{trans}} R \ln(\delta x/1660) \quad (9)$$

where δx are the root-mean-square (rms) amplitudes of translational movements in principle directions, n_{trans} is the number of translational degrees of freedom, and $V_0 = 1660$ Å^3 is the volume available to a single particle in a solution at standard state. The translational component of entropy is the only entropy component that explicitly depends on concentration.

Similarly the change in the rotational entropy, $S_{\mathrm{rot}}^{\mathrm{ideal}} = R \ln[(8\pi^{7/2}\, \mathrm{e}^{3/2}/\sigma)(I_{\mathrm{A}}I_{\mathrm{B}}I_{\mathrm{C}})^{1/2}(2k_{\mathrm{B}}T/h^2)^{3/2}]$, is estimated by

$$\Delta S_{\mathrm{rot}} \approx R \ln(\delta\alpha_1 \delta\alpha_2 \delta\alpha_3/(8\pi^2)) \approx n_{\mathrm{rot}} R \ln(\delta\alpha/4.3) \quad (10)$$

where σ is the symmetry number (it is equal to 1 for asymmetric molecules and to 2 for dimers, etc.), $I_{\mathrm{A,B,C}}$ are the principal moments of inertia, n_{rot} is the number of rotational degrees of freedom, and $\delta\alpha_{1,2,3}$ are the rms amplitudes (in radians) of oscillatory motions about principal directions, as determined from the motions of the molecules in the crystal structures.

Internal atomic vibrations contribute to the vibrational entropy of a molecule. We estimated them for cyt c_2 from a normal-mode analysis using the elNémo webserver.[92] Generally only a few normal modes contribute substantially.

As the binding free energy is the difference between the free energy of the complex and each of the proteins, the geometric components cancel when data from a single trajectory are used for the analysis. In the single-trajectory application of the MM-PBSA method[56−58] with statistical thermodynamic calculations of entropy,[63] the binding free energy can be expressed as

$$\Delta G_{\mathrm{bind}} = \Delta E_{\mathrm{Coul}} + \Delta E_{\mathrm{VdW}} + \Delta G_{\mathrm{solv}}^{\mathrm{polar}} + \Delta G_{\mathrm{solv}}^{\mathrm{nonpolar}} - T(\Delta S_{\mathrm{trans}} + \Delta S_{\mathrm{rot}} + \Delta S_{\mathrm{vib}}) \quad (11)$$

While the entropic contributions are calculated from a single crystal structure (PDB ID 1L9B), all of the other energetic terms reported are averages calculated over the last nanosecond of the 4 ns equilibration runs.

2.6. Methods: Evolutionary Analysis. *2.6.1. Sequence and Structural Data.* The c-cyt representative structures were extracted from either the Protein Data Bank (PDB)[65] or the SCOP database's family of monodomain cyt c's.[38] To find additional sequences for homology studies, sequence profiles from structural representatives were used in a BLAST search[66] with an *E*-value cutoff of 10^{-5} over the Swiss-Prot database.[67] Sequence and structure data for the RC and cyt c_1 were collected in a similar manner using structure profiles and based on the structures from species of interest.

2.6.2. Multiple Alignments. Structural alignments of cyt c_2 were calculated using STAMP[68] as implemented in VMD 1.8.4.[46,69] STAMP is a multiple structural alignment program that employs a combination of a dynamic programming algorithm with least-squares minimization of the pair distances between α carbon atoms to find optimal rotations for each structure. The alignments were executed using the following STAMP parameters: number of passes (npass) 2, similarity (scanscore) 6, and comparison of residues (scanslide) 5. The structure-based alignments were checked for fluxional loop regions as described in ref 39, and it has been found that the core of all used structures superimposed well upon the *Rb. sphaeroides* reference structure.[70]

Sequence alignments of the RC subunits and cyt c_1 were performed with the progressive multiple sequence alignment procedure as implemented in CLUSTALW[71] as implemented in VMD.[46,69,97] Sequence QR factorization[39] was applied with a 90% sequence identity threshold to reduce redundancy, while retaining specific sequences of interest. The full alignments for the RC subunits and cyt c_1 are given in the Supporting Information. In the text only a representative set is given that includes proteins used in experimental studies. The progressive sequence alignment method does not always ensure the optimal alignment between distant homologues, and extraneous gaps can be introduced by the algorithm. As structural alignments are generally more reliable than sequence alignments,[72,39] sequence alignments were manually checked between cyt c_2 proteins in different classes and adjusted when necessary against the structural alignment.

2.6.3. Phylogenetic Analysis. The structure-based phylogenetic trees were drawn by employing either the neighbor-joining program in Phylip version 3.6[73] or the unweighted pair group method with arithmetic averages (UPGMA, dendrogram)[74] as implemented in MATLAB 6.5 (Mathworks, Natick, MA). The measure of distance was $1 - Q_{\mathrm{H}}$, where Q_{H} is a structural similarity measure developed by O'Donoghue and Luthey-Schulten[37] to account for both the aligned and the gapped regions. Sequence-based phylogenetic trees for multiple sequence alignments were constructed as distance-based neighbor-joining trees with the percent accepted mutation (PAM) substitution matrix metric using Phylip. TreeView[75] and the MATLAB phylogenetic tree tool were used to draw trees.

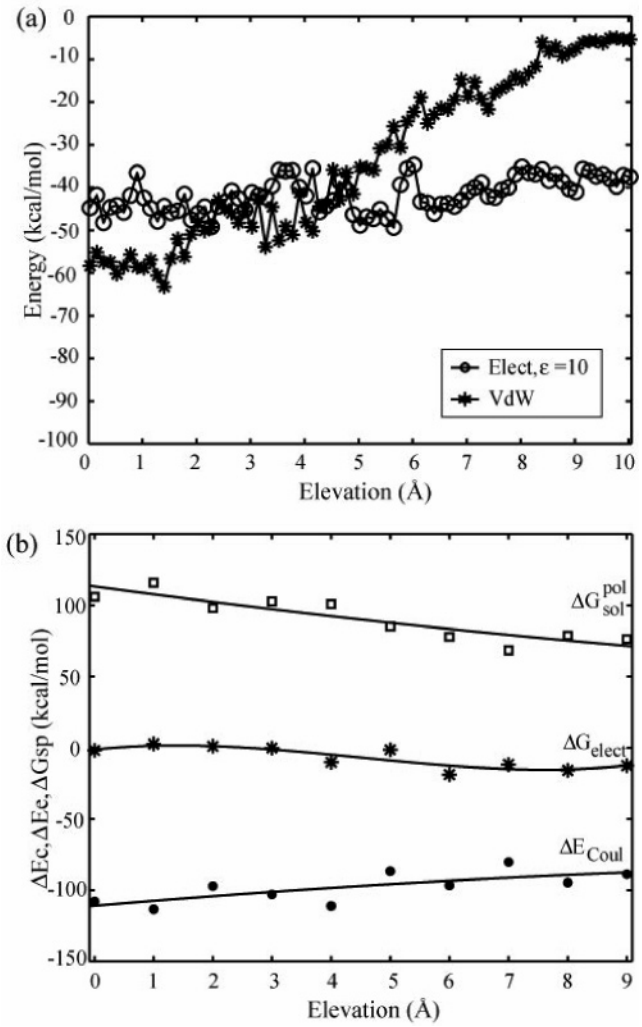

Figure 4. Nonbonded interaction energies during the SMD elevation of cyt c_2^{ox}. (a) VdW and electrostatic contributions calculated using CHARMM27 in MDenergy. The electrostatic contribution was scaled with a dielectric constant of 10. (b) Polar solvation, Coulombic, and total electrostatic energies calculated with DelPhi with $\epsilon_{in} = 4$, $\epsilon_{out} = 80$, and PARSE charges.

3. Results and Discussion

3.1. Possible Exit Pathways for Cytochrome c_2^{ox}. Figure 3 shows a possible exit pathway for cyt c_2^{ox} based on the VdW interaction energy between the RC and cyt c_2^{ox}. As previously described[20] there is a broad VdW minimum extending more than 2 Å in the *x*- and *y*-directions from the initially minimized RC−cyt c_2^{ox} docked position at the origin. In the current study, we sampled the VdW interaction energy in a 40 Å by 40 Å square centered at the docked position of cyt c_2^{ox} and the RC with cyt c_2^{ox} elevated to 1, 2, and 3 Å. The energies at the elevation of 3 Å revealed a broad unobstructed pathway across the L subunit extending toward the surrounding membrane. Sampling across the RC at lower elevations of 1 and 2 Å did not reveal a clear pathway of dissociation (data not shown).

The electrostatic interactions between the RC and cyt c_2^{ox} as a function of elevation from its docked position at the origin are shown in Figure 4. Similar to the energy landscape analysis in the previous study, an SMD simulation with explicit waters was used to pull cyt c_2 to a height 10 Å from the docked position at a rate of 5 Å/ns. The VdW contribution displays a relatively slow reduction, with values fluctuating around −55 kcal/mol until the distance between the centers of mass of the RC and cyt c_2 was increased to 2.5−3 Å. At this point important hydrophobic contacts start to break, and the VdW contribution steadily declines.

To recover the electrostatic contribution of the protein−protein interaction, the Coulomb energies calculated between the RC and cyt c_2 during the SMD trajectory were scaled with a dielectric constant of 10, as in the previous reports.[20,76] It has been shown by McCammon and co-workers[77] and Zhou and Schulten[78] that at least four hydration layers of water molecules are required to transition to bulk properties at a separation of ±25 Å. By the elevation level of 6−7 Å the transient salt bridges between AspL261 and LysC97/LysC99 and the cation−π interaction between TyrM295 and ArgC32 are no longer present, and the electrostatic contributions decrease by 5−10 kcal/mol with respect to the values at the beginning of the pull. Interestingly, at the height of 2.5−3 Å VdW energy becomes comparable to the electrostatic energy and then continues to decrease at a faster rate, showing the importance of VdW contribution to undocking.

The electrostatic contribution to binding free energy for each conformation in the SMD trajectory can also be estimated (see the Methods section). The Coulombic contribution as shown in Figure 4b has values close to −100 kcal/mol for the initial 4 Å of elevation starting from 0 Å separation. The polar contribution to solvation free energy ΔG_{sol}^{pol} steadily decreases with increasing separation. Notably the total electrostatic binding free energy, which is the sum of the above two components, is slightly repulsive at low elevations and only becomes slightly attractive when cyt c_2^{ox} is elevated higher than 3 Å above its docked position. The behavior of the electrostatic contribution for the oxidized system agrees well with the calculations by Onuchic and co-workers[76] performed on the reduced system that were not equilibrated in the presence of explicit water molecules. The repulsive nature of the electrostatic contribution ΔG_{elect} at low elevations supports cyt c_2^{ox} dissociation along either the vertical or the horizontal pathways and clearly points out the importance of the van der Waals interactions in stabilizing the bound complex.

3.1.1. Exploring the Proposed Undocking Pathway. It had already been suggested that to maintain efficient turnover cyt c_2^{ox} should leave in the direction opposite to the approach of cyt c_2^{red}.[23,30,31] Recent experimental and computational studies have indicated that cyt c_2^{red} forms an initial encounter complex with the RC above the M subunit.[7,30] As the encounter complexes can already be formed when cyt c_2^{red} is as much as 10 Å away from the docking site in the X-ray structure[30] on the M subunit, an exit path along the L subunit is not unexpected. The location of the energy barrier between the encounter complex and the docked complex is consistent with our scanning results in Figure 3.

During the dissociation study the applied force varies along the horizontal displacement at 3 Å elevation as shown in Figure 5a. It reaches a maximum value fluctuating between 600 and 1000 pN. In this region the electrostatic and VdW energies are approximately constant. As cyt c_2^{ox} is pulled along the RC surface transient salt bridges are continuously being formed and broken (Figure 5c), starting initially with two contacts AspL261-O_δ−LysC97-N_ζ and AspL257-O_δ−LysC99-N_ζ which fluctuate in a similar fashion as observed during the equilibration.[20] The rapid decrease in force at approximately 15 Å separation can be directly attributed to the irreversible breakage of the VdW contacts between LeuM191−ThrC17 and ValM192−PheC102, as seen in the docked state (Figure 5d). After the breakage of the VdW contacts the force increases again due to new

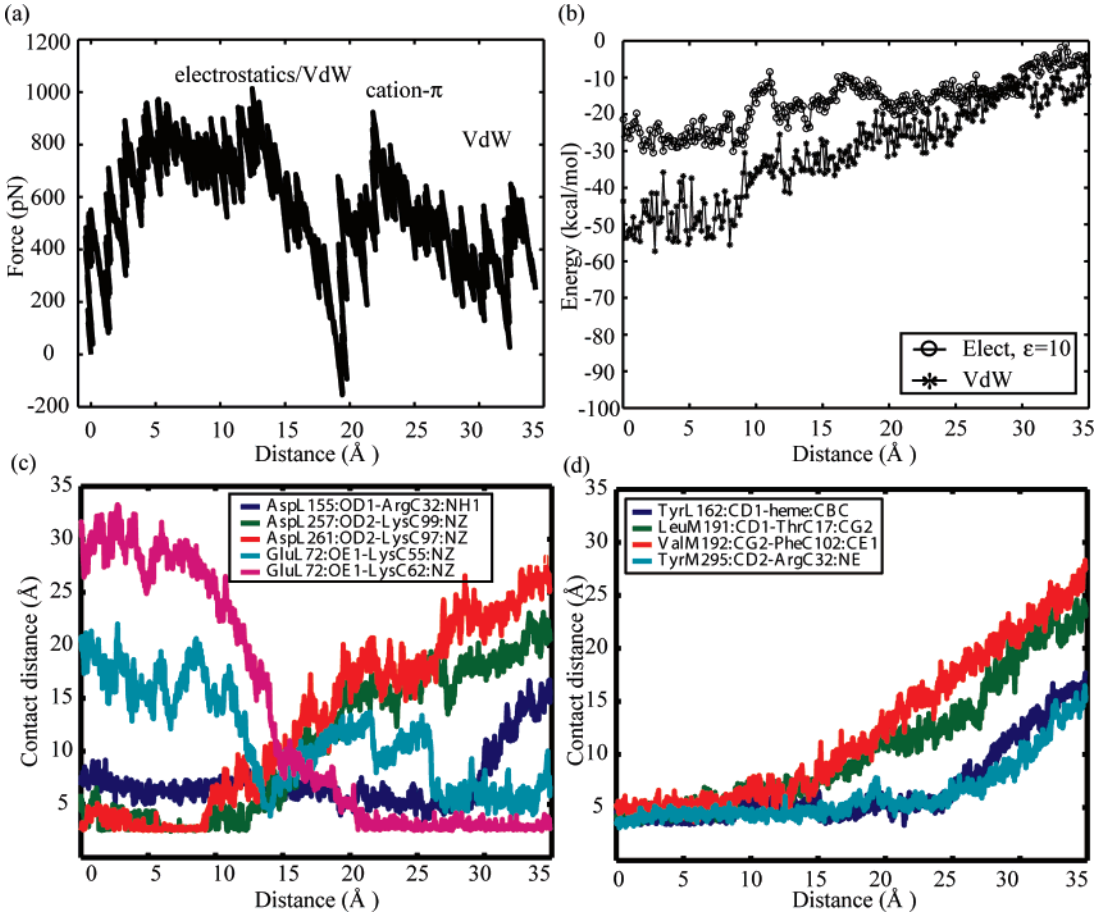

Figure 5. (a) SMD force plot as cyt c_2 (elevated to 3 Å) is pulled horizontally along the L subunit. The rapid decrease in force starting at approximately 22 Å separation can be directly attributed to the irreversible breakage of the VdW contacts and cation$-\pi$ pair. (b) Nonbonded interaction energies during the SMD pull. The favorable VdW interaction energy decreases significantly when the hydrophobic interface is broken. In contrast the electrostatic energy contribution decreases by approximately 10 kcal/mol as weak electrostatic contacts establish themselves during the whole course of the SMD simulation. (c) Electrostatic contacts and (d) VdW and cation$-\pi$ contacts during the horizontal cyt c_2 pull along the RC surface. VdW contacts remain stable until ~15 Å separation and then become irreversibly broken. In contrast electrostatic contacts can be broken and newly formed during the whole course of the SMD simulation.

electrostatic interactions between GluL72−LysC55 and GluL72−LysC62. The final decrease of the force corresponds to the breakage of the remaining VdW pairs and the cation$-\pi$ interaction between ArgC32 and TyrM295. The nonbonded energy contributions in Figure 5b show the significance of short-range VdW interactions during the initial pulling phase. In the unperturbed complex the VdW contribution for the interaction energy is approximately 30 kcal/mol larger than the electrostatic contribution. The VdW energy contribution fluctuating around −50 kcal/mol becomes less favorable by approximately 20 kcal/mol when the short-range docking domain is broken. This energy value decreases continuously to −10 kcal/mol during the course of the SMD run, at which point cyt c_2^{ox} has reached the edge of the RC. Overall, the VdW energy decreases by 40 kcal/mol between the docked cyt c_2^{ox} and the end of the SMD simulation, whereas the strongly fluctuating electrostatic energy contribution decreases by only 10 kcal/mol, when scaled with the dielectric constant of 10. When the force is applied, contacts TyrL162−heme and TyrM295−ArgC32 remain stable for the longest period of time (Figure 5d). As cyt c_2 is a globular protein its center of mass can rotate toward the pulling direction before breaking these nonpolar contacts.

The cation$-\pi$ interaction formed through direct contact between a charged cyt c_2 residue (ArgC32) and an aromatic RC residue (TyrM295) establishes the most stable contact observed in our simulation. This contact is only broken after cyt c_2^{ox} has been pulled approximately 16 Å away from its initial docking position. The largest local conformational change at the periplasmic surface of cyt c_2 is attributed to the formation of the cation$-\pi$ interaction pair.[7] In a very recent experimental study by Okamura and co-workers[21] the cation$-\pi$ interaction pair has been subjected to site-directed mutagenesis. The 80-fold increase in the dissociation constant observed upon mutation of ArgC32 to a neutral alanine confirms the stabilizing role of the electrostatic interaction. The authors attributed the much smaller change in the dissociation constant for the ArgC32Lys mutation to the lack of cation$-\pi$ pair formation, but our evolutionary analysis presented in Figure 9a reveals a high replacement rate of arginine by lysine in purple bacteria, suggesting formation of an important electrostatic, possibly cation$-\pi$, interaction.

Interestingly, when cyt c_2^{ox} was pulled above the membrane without the RC, the characteristics of the SMD force and of the nonbonded interactions are remarkably different. After initial growth, the SMD force plateaus, revealing the homogeneity of the membrane (Figure 6a). The nonbonded energy contributions confirm that no major variations are present. A slight reduction of the electrostatic energy is due to the relaxation of the side chains as cyt c_2^{ox} is being pulled. At an elevation of 3 Å the electrostatic interaction between cyt c_2^{ox} and the membrane is slightly more favorable than its interaction with the RC (Figure 6b). The VdW interaction energies are comparable to the values

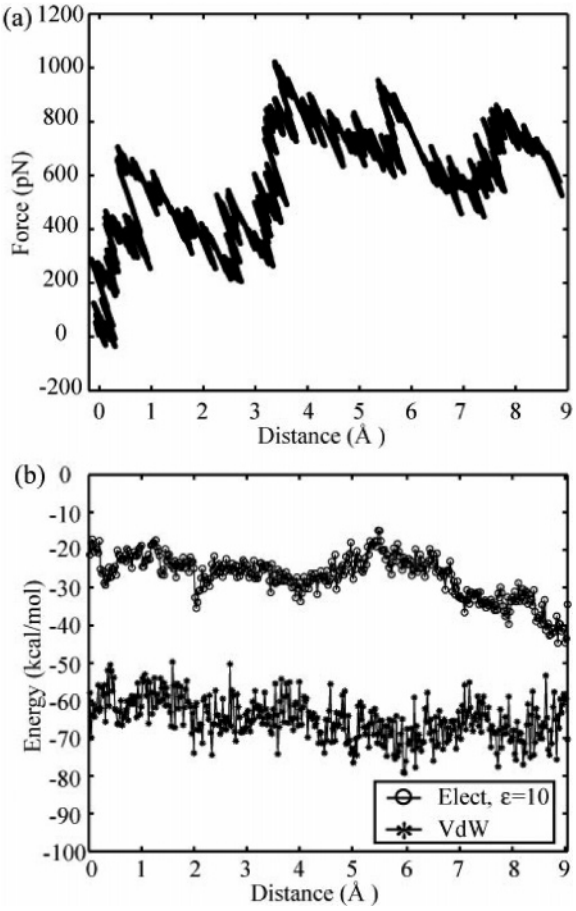

Figure 6. (a) SMD force plot as cyt c_2 (elevated to 3 Å) is pulled horizontally along a membrane. The initial increase in the force is due to initialization of the constant velocity protocol. (b) Nonbonded interaction energies during the SMD pull over the membrane. VdW interaction energies are nearly constant. The electrostatic energy contribution becomes more favorable as residues relax.

observed in the docked position, suggesting that a somewhat higher elevation would be favorable for its transport to the next docking partner. The docking site for cyt c_2^{ox} on the cyt bc_1 complex is 10 Å above the membrane, and from our studies it is clear that by an elevation of 10 Å the VdW interaction energies with the membrane are nearly zero.

3.2. Evolutionary Studies. Evolutionary analysis identifies both specific structural and sequence features that have developed for cyt c_2 to interact with its partners, the RC and cyt bc_1. These features can also be used to analyze the SMD simulation results along the suggested exit pathway and to predict binding characteristics for the complexes from various species.

3.2.1. Structure and Sequence Phylogenies of Type c Cytochromes. As structure is more conserved than sequence,[36,37,80] we first examine the evolution of the structure and identify those specific features important for docking to either RC or cyt bc_1. Structures of water-soluble, type c cytochromes are generally divided into three classes according to their sizes: small (cyt $c_{6/5}$, S), medium (cyt c and cyt c_2, M), and large (cyt c_2, L).[17] The phylogeny shown in Figure 7a depicts the evolution of the structures of water-soluble, type c cytochromes based on a structural metric, Q_H, recently developed by O'Donoghue and Luthey-Schulten,[37] which takes into the account gaps in structure-based alignments. As can be seen, the phylogeny displays two major branches, one representing Cyanobacteria, the chloroplasts, and γ-Proteobacteria and the other containing the α-Proteobacteria and mitochondria. The α-Proteobacteria/mitochondria branch is subdivided into two subbranches. The first subbranch contains mitochondrial cyt c along with the so-called medium α-Proteobacteria cyt c_2, and the other subbranch is comprised solely of the large α-Proteobacteria cyt c_2.

Cytochrome c structures from mitochondria are structurally closely related to cyt c_2 structures from *Rhodopila* (*Rph.*) *globiformis* and *Blastochloris* (*Blc.*) *viridis*.[17,18] In these photosynthetic bacteria a tetraheme complex docks directly to the RC, and it is still not known where cyt c_2 binds to the RC complex. The large α-Proteobacteria/mitochondria subbranch is comprised solely of purple bacteria cyt c_2 and is congruent with the sequence phylogenetic tree reported by Woese and coworkers.[18,35,81] *Rb. sphaeroides* cyt c_2 is located in this lower subbranch with another five representatives of the α-Proteobacteria, all of which have additional structural insertions characteristic of the large cytochromes as described in ref 17 and as seen in the alignments in Figure 9a. Sequence identity of *Rb. sphaeroides* and *Rb. capsulatus* cyt c_2 is 50%, and they have a structural overlap with a root-mean-square deviation (rmsd) of 1.44 Å, which corresponds to Q_H of 0.71. In purple bacteria, cyt c_2 can either dock directly to the periplasmic site of the photosynthetic RC or dock to the so-called cytochrome subunit, which has been described for the RC X-ray structure of *Blc. viridis*.[6,82]

3.2.2. Structure Conservation. The core structures of both water-soluble and membrane-bound type c cytochromes are well-conserved.[96] In Figure 8 the structure of cyt c_2 of *Rb. sphaeroides* is color-coded according to the structure conservation (Q_H) with respect to the 19 water-soluble proteins presented in Figure 7. The most conserved elements are the N- and C-terminal helices and the 60s helix labeled in the alignments. Insertions characteristic of the large cyt c_2 appear as red on either side of the protein. Interfacial residues within 5 Å of each other are given in licorice on cyt c_2 and in VdW surface representation on the RC. Most of them, with the exception of ArgC32, which is located in an insertion specific to the large cytochromes, are located in the structurally conserved regions. Interestingly, the RC interaction residues, color-coded by sequence similarity, are also mostly conserved (Figure 8).

3.2.3. Sequence Conservation. The multiple sequence alignment of nine cyt c sequences representing the three size classes is shown in Figure 9a. The alignment was initially constructed using structures of cyt c from the S, M, and L classes and then corrected for deviation of conformations according to sequences of cyt c from the M and L classes. The multiple sequence alignment reveals a relatively small number of conserved residues (close to 10% identity between cytochromes of the S and L classes). The conserved residues can be divided into four classes: The first is the heme-binding signature motif CXYCH (X and Y are any residue) plus methionine found at positions 15−19 and 100 in *Rb. sphaeroides*, respectively. The cysteine residues, CysC15 and CysC18, covalently attach the heme, and HisC19 and MetC100 are heme ligands.[38] MetC100 is located on the loop before the C-terminal helix. Second, the residues located on the N- and C-terminal α-helices have been found to be important for folding as they establish helix−helix contact: conserved glycine (position 8), which makes contact with the backbone at the residues aligned with IleC113, and mostly aromatic residues aligned with PheC12, which make contact with the residues aligned with TyrC116.[83] Third, several highly conserved residues are in direct contact with the heme prosthetic group, such as ArgC46 (Figure 9a). And finally are the residues that we hypothesize are conserved due to their importance in protein−protein interaction with either RC or cyt bc_1. The

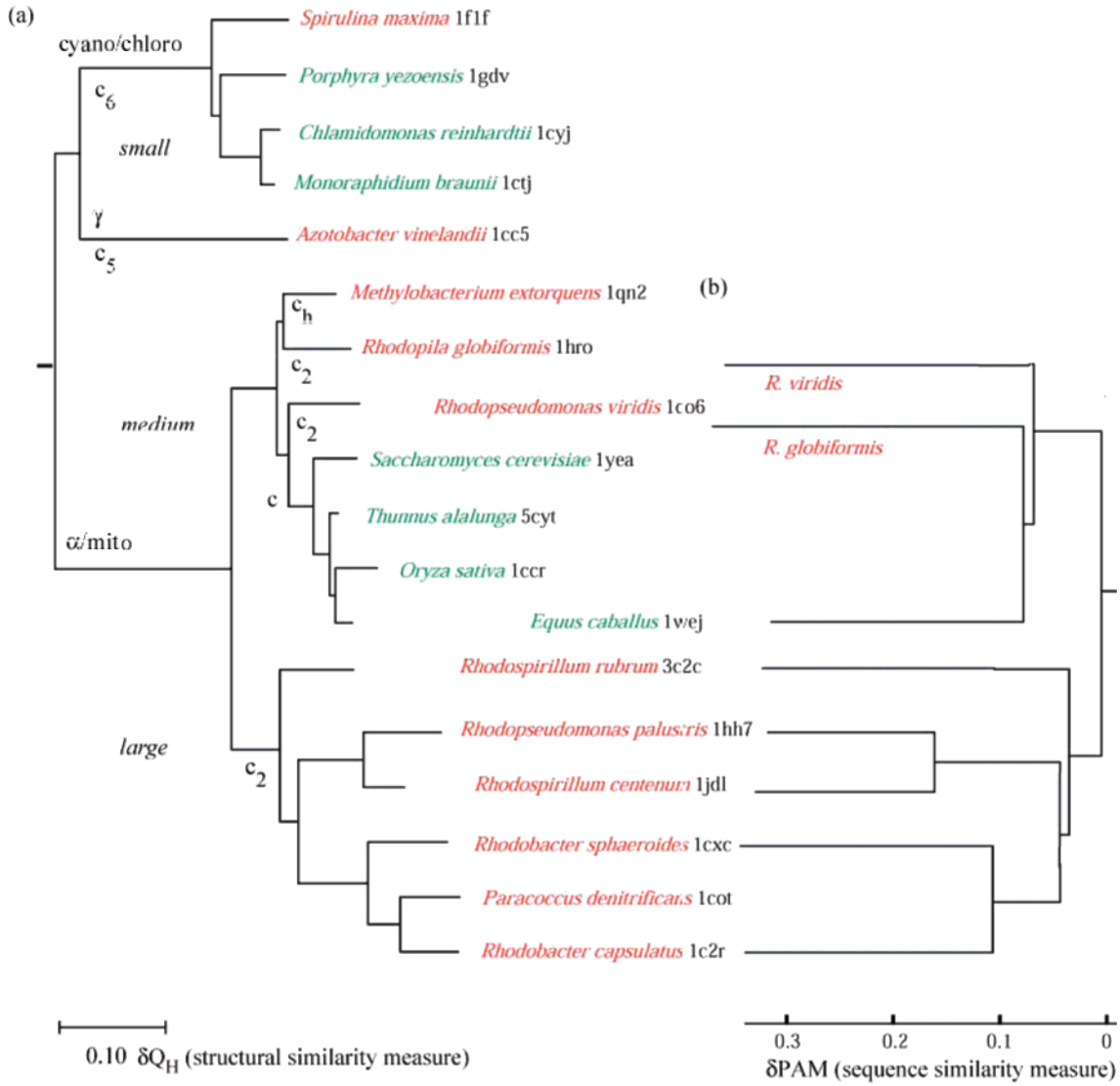

Figure 7. Phylogenetic relationships of cytochromes. (a) Phylogeny based on structures from chloroplasts (c_6), mitochondria (c), and α-Proteobacteria (c_2). The structure of cyt c_{552} from *Thermus thermophilus* was used as an outgroup (not shown). Two main branches are formed, one from cyanobacteria and chloroplasts and the other from α-Proteobacteria and mitochondria. Eukaryotes are presented in green, and bacteria are in red. (b) Sequence-derived phylogeny based on cytochrome sequences from purple bacteria and mitochondria. The sequence of cyt c_5 of a γ-Proteobacteria, *Azotobacter vinelandii*, was used as an outgroup (not shown). This comparison confirms that phylogeny can be obtained from both sequence and structural alignments for the α-Proteobacteria and mitochondria groups.

conservation of these residues is not strict across all species as cyt c has various docking partners even in closely related species. In the case of *Rb. sphaeroides*, *R. rubrum*, and *Rb. capsulatus* where cyt c_2 docks to the periplasmic site and transfers its electrons directly to the special pair, the conservation of lysine residues at the docking site is striking (shown in boxes). However, none of the lysine residues at the cyt c_2 docking site exhibits complete conservation across all species. The cluster of lysines involved in contacts in the docked complex in cyt c_2 of *Rb. sphaeroides* containing LysC97, LysC99, and LysC103 is conserved and establishes weak electrostatic interactions with the corresponding negatively charged RC residues AspL261, AspL257, and AspM184/GluM95. The conservation becomes more apparent when one takes into account that lysine residues located close to this cyt c_2 sequence position in the other species are separated very often by only one position and still might form contacts in the docked structures. Small cytochromes lack many of the lysine residues in the region from LysC97 to LysC106. As the latter is located on a loop region of the cytochromes, we propose that the addition of the charged residues was guided by a functional need. As shown in Figure 9b, AspM184, GluM95, and AspL261 are among the most conserved periplasmic RC residues. These residues have been also shown to be important for RC−cyt c_2 docking in mutational studies.[23,29,31] The conserved LysC103 and LysC106 are predicted to interact with the cyt bc_1 complex based on our model discussed below and would be good candidates for further mutagenesis experiments.

The conservation of residues forming the hydrophobic electron-transfer domain in both purple bacteria and mitochondria is well established in the literature.[2,4] As can be seen in Figure 11 the central contact area on the RC interface in *Rb. sphaeroides* is formed by the conserved residues TyrL162 and LeuM191, which are both conserved across all purple bacteria (see also the Supporting Information), and ValM192. The corresponding residue partners on cyt c_2 are mainly the heme edge and the conserved PheC102. From the sequence alignment presented in Figure 9b, it is evident that ValM192 is not conserved, and experimental data confirm that mutation of this residue only marginally changes the dissociation constant and

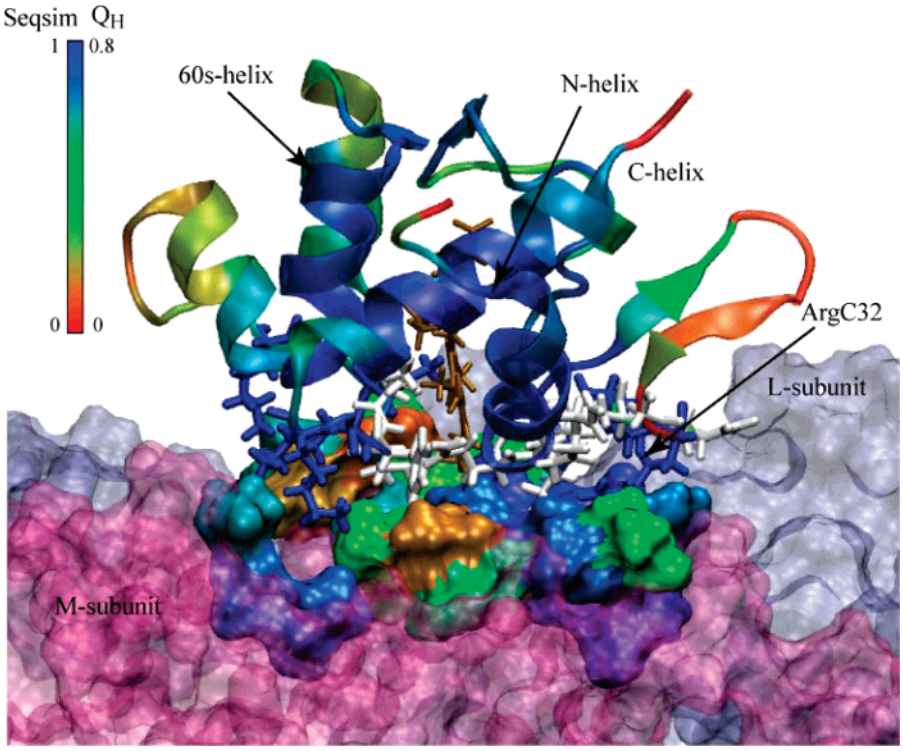

Figure 8. Sequence and structure conservation of the RC and cyt c_2 in *Rb. sphaeroides*. Residues within 5 Å are shown in licorice (cyt c_2) and in VdW (RC) and color-coded by type (positively charged (blue) and noncharged (white)) and by sequence similarity (BLOSUM50), respectively. Sequence similarity of the RC subunits is calculated based on alignments of a larger set of sequences presented in the Supporting Information. Cytochrome c_2 is color-coded by structural similarity (Q_H) over the set of structures presented in Figure 7a.

the rate of electron transfer.[22] Thus this type of evolutionary analysis can be used to gauge the importance of a given residue for association and dissociation of protein complexes.

This analysis in combination with MD studies allows us to suggest residues that can be important for the dissociation process. Charged residues on the L subunit of the RC, AspL155 and GluL72, which we propose to be important for undocking of cyt c_2, are conserved, especially if one takes into account other neighboring species of α-Proteobacteria. Additionally, the cation$-\pi$ interaction has been established as an important contributor to protein−protein interfaces.[84] In this study we also observe the high stability of this interaction. As shown in Figure 9 the ArgC32−TyrM295 cation$-\pi$ contact is a highly conserved contact. From full alignment the cation$-\pi$ pair of either ArgC−TyrM or an equivalent LysC−PheM pair is a conserved motif in all purple bacteria with the exception of *R. rubrum*. It was shown[35] that *R. rubrum* branches off far from the common evolutionary origin. This is an indication that the cation$-\pi$ interaction is an ancient RC−cyt c_2 contact that can be lost during the course of evolution in purple bacteria.

3.2.4. Cytochrome: Various Roles. Cytochrome c docking partners in mitochondria are cyt c oxidase and the cytochrome bc_1 complex. All complexes of the mitochondrial respiratory chain have homologues in purple bacteria.[4] When cyt c_2 docks to the cytochrome bc_1 complex, it binds to the cyt c_1 domain. We created a model of a docked bacterial cyt c_2−cyt bc_1 complex by overlapping the structure of the cyt c−cyt bc_1 complex of yeast (PDB ID 1KYO[8]) with the individual structures of the cytochrome bc_1 complex (PDB ID 1ZRT[9]) and cyt c_2 from *Rb. capsulatus* (PDB ID 1C2R). The model, shown in Figure 10a, allowed us to characterize the interface between cyt c_2 and the cytochrome bc_1 complex in the purple bacteria. Even though the interface area is smaller than that in the case of the RC−cyt c_2 complex, cyt c_2 uses many of the same residues to dock to both the cytochrome bc_1 complex and the RC. The docking to cyt c_1 should be viewed more like a dimerization event in which the exposed edges of the hemes are coplanar but the structural cores are rotated by approximately 120° with respect to each other. The latter is evident from the fact that only one region of cyt c_2 (the loop between the C-terminal helix and the helix 4, blue) is aligned with the structurally equivalent motif of cyt c_1 (Figure 10c). Additionally, an interesting evolutionary development came to light when we studied the changes between the cyt c_2 (large), cyt c (medium), and cyt c_1 structures (Figures 10b and 10c). The structural alignment of the *Rb. capsulatus* and yeast proteins reveals the neutralization of the positively charged residues present in the docking domains of cyt c_2 and cyt c and the addition of extensive negatively charged loops, assumed to guide cyt c_2 to the cyt bc_1 complex. Once a crystallographic structure of the complex of cyt c docked to cytochrome oxidase becomes available, the evolutionary picture of cyt c_2 complexes will be complete. For example, the function of the large insertion surrounding helix 2 is still not clear. There is some indication that this helix modifies the redox potential of cyt c and cyt c_2 by protecting the propionic acids on the heme from exposure to solvent,[91] but its effect on the docking of cytochrome to its partners, if any, is not known.

3.2.5. Variance of Docking Interfaces. It is interesting to ask if the evolutionary studies along with the mutational data can be used to predict the change in binding affinity between the RC and the large cyt c_2 for closely related organisms. From the conservation observed in the sequence alignments in Figure 9, it is clear that the RC−cyt c_2 complex from *Rb. capsulatus* should have the same kinetic behavior as *Rb. sphaeroides*, which has been confirmed experimentally.[31] The situation for *R. rubrum* is somewhat unclear. In Figure 11 we compared the docking surfaces of the *Rb. sphaeroides* RC−cyt c_2 complex[7]

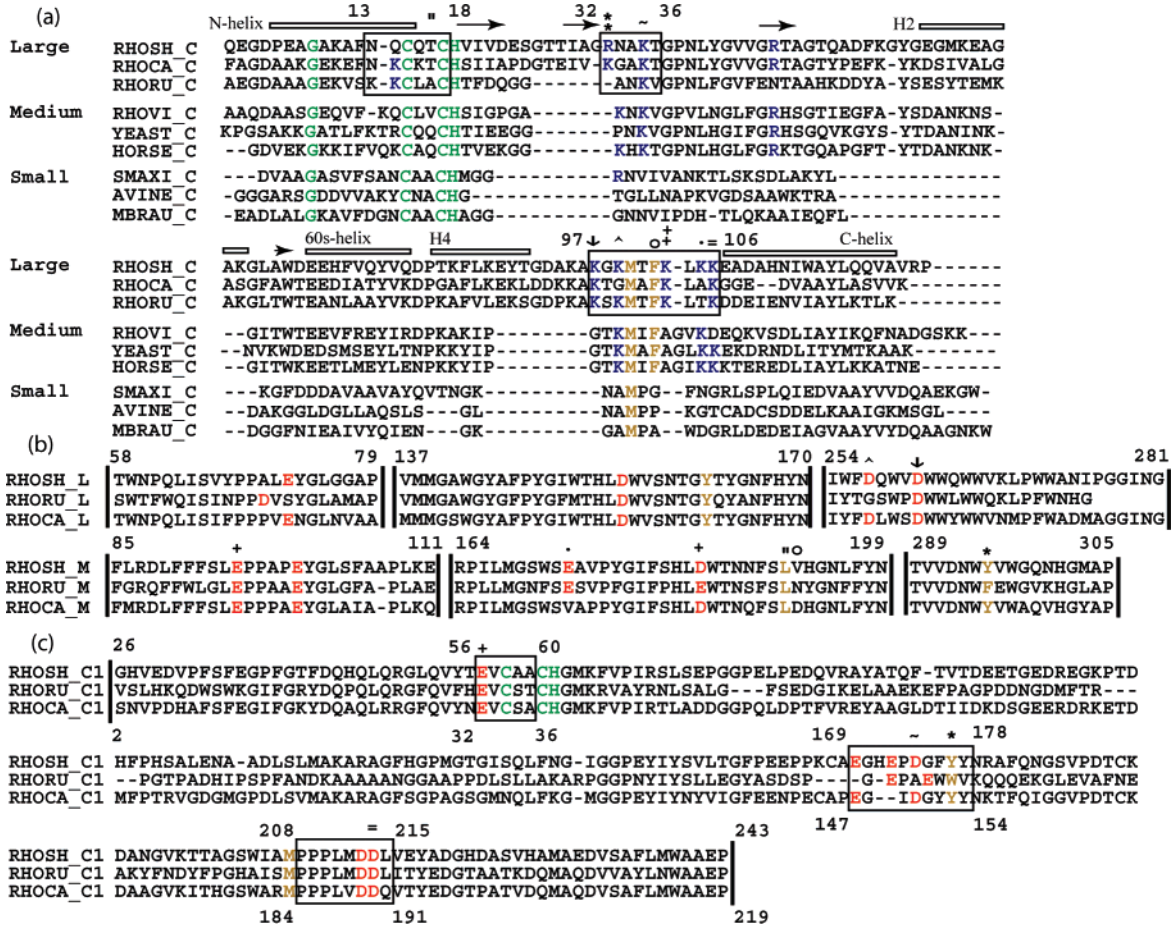

Figure 9. Multiple sequence alignments of (a) large, medium, and small cytochromes, (b) L and M subunits of the RC, and (c) cyt c_1 of the cyt bc_1 complex. The cytochrome docking domains are indicated by boxes (a and c). The corresponding docking domains of the RC and cyt c_1 are depicted in parts b and c. The RC sequence between the periplasmic domains and the anchor domains of cyt c_1 has been omitted for clarity. Residues important for cyt c_2 interactions are marked and are described in the text. Amino acids are numbered with respect to their positions in *R. sphaeroides*. Color code of the residues: nonpolar, brown; polar, green; positively charged, blue; negatively charged, red. The following organisms have been used: *R. sphaeroides* (RHOSH), *R. rubrum* (RHORU), *Rb. capsulatus* (RHOCA), *B. viridis* (RHOVI), *S. cerevisiae* (YEAST), *E. caballus* (HORSE), *S. maxima* (SMAXI), *A. vinelandii* (AVINE), and *M. braunii* (MBRAU).

with those of the cyt c_2 X-ray structure from *R. rubrum*[85] together with an RC model generated for *R. rubrum*.[45] Although most of the important interface residues are conserved, there are three significant differences. First, *R. rubrum* cyt c_2 does not contain the protruding loop with ArgC32; therefore the specific cation$-\pi$ interaction pair could not be established in *R. rubrum* RC$-$cyt c_2 complexes. Second, *R. rubrum* is missing the buried AspM88 of *Rb. sphaeroides*, which causes a 55-fold decrease in binding affinity after mutation to lysine.[23] Third, AspM184, the residue located in the region of strongest electrostatic interactions investigated in single and double mutational studies,[23,31] is replaced in *R. rubrum* by a homologous glutamic acid with a longer side chain. These distinct differences between the two docking interfaces from *Rb. sphaeroides* and *R. rubrum* should have direct implications on the docking affinity and specificity of the respective RC$-$cyt c_2 complexes.

In a short preliminary molecular dynamics simulation on the *R. rubrum* system,[45] we observed a docked complex similar to the X-ray structure of in *Rb. sphaeroides*. During the simulations, salt bridges formed between GluM182 (AspM184 in *Rb. sphaeroides* numbering) and two lysine cyt c_2 residues, LysC75 and LysC84. The latter corresponds to LysC103 in cyt c_2 of *Rb. sphaeroides* and the former corresponds to ThrC84, which is a neighbor of LysC85. The salt bridge with LysC75 is stable over the course of 200 ps of equilibration and was already formed after minimization of the initially docked RC$-$cyt c_2 complex. The salt bridge with LysC94 was formed after 25 ps and was stable until the end of the 200 ps simulation. These results are comparable to those reported above in the equilibration runs in the *Rb. sphaeroides* complex that revealed the presence of at least two closely positioned pairs of charged residues capable of forming transient salt bridges. These residues are AspL261, AspL257, LysC97, and LysC99. Additionally, the multiple sequence alignment of the cytochrome sequences presented in Figure 9 shows that two of the three docking domains of cyt c_2 of *Rb. sphaeroides* have a larger number of charged residues. As the loss of the cation$-\pi$ interaction would increase the dissociation rate and the change in the interface charge would decrease the association rate, we expect a larger dissociation constant for cyt c_2 binding to the RC in *R. rubrum*. The latest, to our knowledge, experimental studies involving the binding of cyt c_2 of *R. rubrum*, performed in the late 1980s,[86] seem to agree with our prediction, but a more detailed study with single site mutations would be needed to clarify the differences in the binding.

When we applied a similar analysis to horse cyt c, the electrostatic map displayed a larger coverage than in the case of cyt c_2 of *R. rubrum* (data not shown). Additionally, the structure-based alignment reveals a large number of positively charged lysine residues in the region of horse cyt c that are aligned to the longest docking region of cyt c_2 of *Rb. sphaeroides*. Finally, ArgC32 of cyt c_2 of *Rb. sphaeroides* is

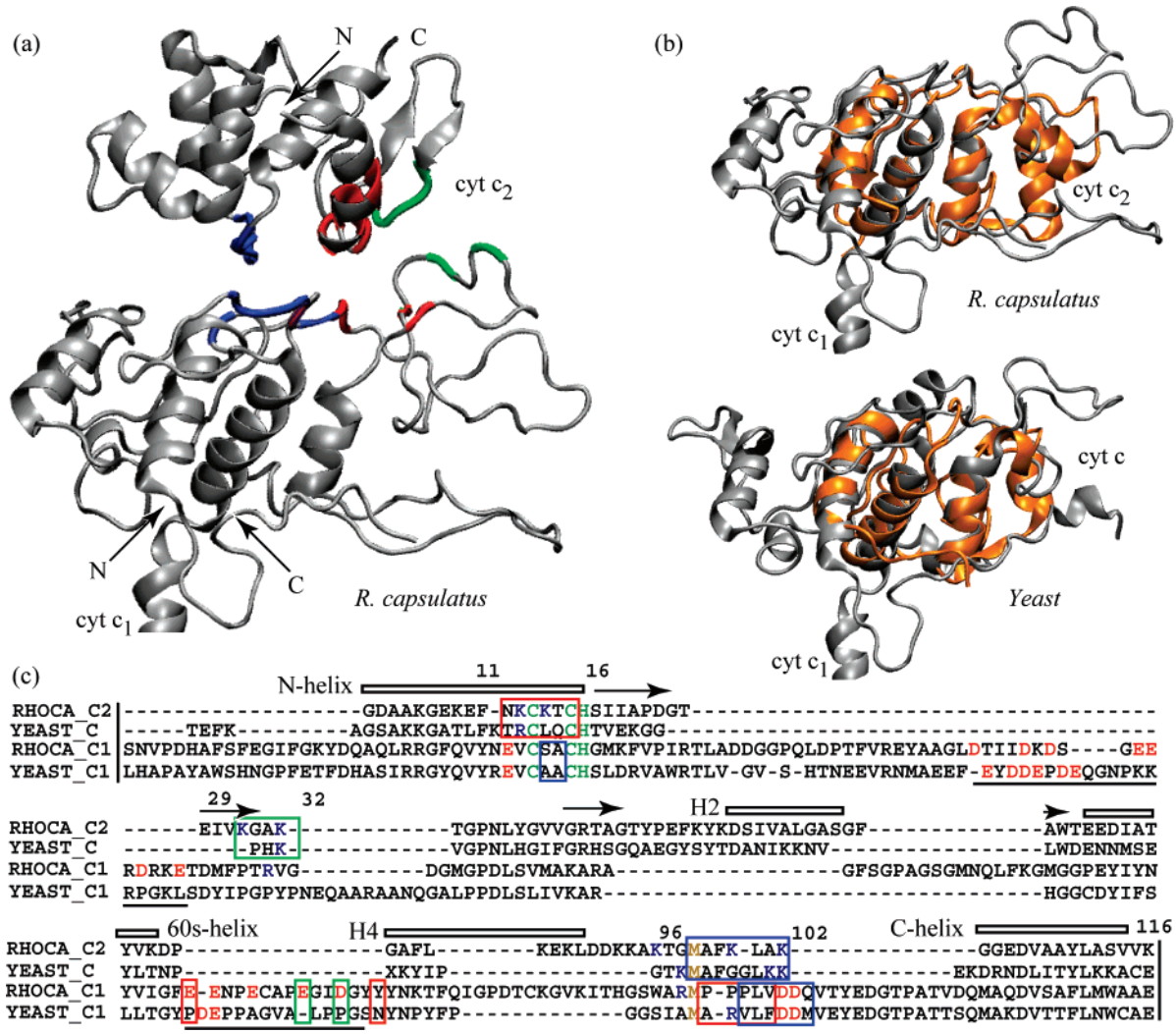

Figure 10. (a) Model of the docked complex of cyt c_2 and cyt bc_1 of *Rb. capsulatus*. (b) Top: Structural alignment of cyt c_2 (gold) and cyt c_1 (gray) of *Rb. capsulatus* (RHOCA), PBD IDs 1C2R and 1ZRT, respectively. Bottom: Structural alignment of cyt c (gold) and cyt c_1 (gray) of *S. cerevisiae* (YEAST), PBD ID 1KYO (c) Corresponding structure-based sequence alignment. The cytochrome docking domains are indicated by color on structures and sequences. Color code of the residues: nonpolar, brown; polar, green; positively charged, blue; negatively charged, red. Secondary structure elements of cyt c_2 are shown as rectangles (α-helices) and arrows (*β*-sheets). Negatively charged loops of cyt c_1 are underlined.

aligned to a lysine of horse cyt c, which may allow for a cation−π interaction. In this case the analysis suggests that cyt c of horse will bind more tightly to the RC of *Rb. sphaeroides* than native cyt c_2 and the dissociation constant will be smaller for the complex with horse cyt c. This conclusion agrees with the experimental data of Larson and Wraight.[87]

3.3. Binding Free Energy of the RC−Cytochrome c_2 Complex. The binding free energies of the RC−cyt c_2 complex for both oxidation states of cytochrome were calculated using the MM-PBSA method. As described in the Methods section an ensemble of structures sampled from full-atom MD equilibration runs are used to determine the energies

$$\Delta G_{\text{bind}} = \Delta E_{\text{Coul}} + \Delta E_{\text{VdW}} + \Delta G_{\text{solv}}^{\text{polar}} + \Delta G_{\text{solv}}^{\text{nonpolar}} - T(\Delta S_{\text{trans}} + \Delta S_{\text{rot}} + \Delta S_{\text{vib}}) \quad (12)$$

In Table 1 we compare the MM-PBSA results to those obtained from a continuum electrostatic model using the program DelPhi. All energy calculations are ensemble averages obtained from protein conformations sampled in the last part of the MD trajectory. The average pairwise Coulomb energies of the RC−cyt c_2 interaction are −398.90 kcal/mol for the oxidized cyt c_2 case and −270.11 kcal/mol for the reduced cyt c_2 case, and the difference simply reflects the effect of the increased positive charge on the cyt c_2 heme group upon oxidation. Similar trends are also seen in the DelPhi results.

As there is a little change in the geometry of the interface upon oxidation of cyt c_2, the average VdW energies are similar. The contributions from the nonpolar component of the solvation energy, $\Delta G_{\text{solv}}^{\text{nonpolar}}$, which is a measure of the energy required to create a cavity for the solute in the solvent, have similar small positive values, indicating that in both cases the interface regions of the complex are slightly more compact than those in the isolated systems.

Formation of the complex buries charges at the interface, making the polar contribution to the solvation free energy positive, $\Delta G_{\text{solv}}^{\text{polar}} \approx 469.63$ kcal/mol and 333.19 kcal/mol for the complexes with cyt c_2^{ox} and cyt c_2^{red}, respectively. As this component reflects the energetic cost of transferring a charged particle to a medium with a higher dielectric constant, the oxidized complex with its larger charge has the higher energy.

On the basis of the effective free energy of binding ΔG_{eff}, which is the sum of the electrostatic, van der Waals, and solvation contributions, the oxidized complex is less stable than the reduced complex by approximately 6 kcal/mol. The largest differences in the free energies arise from the $\Delta G_{\text{solv}}^{\text{polar}}$ terms

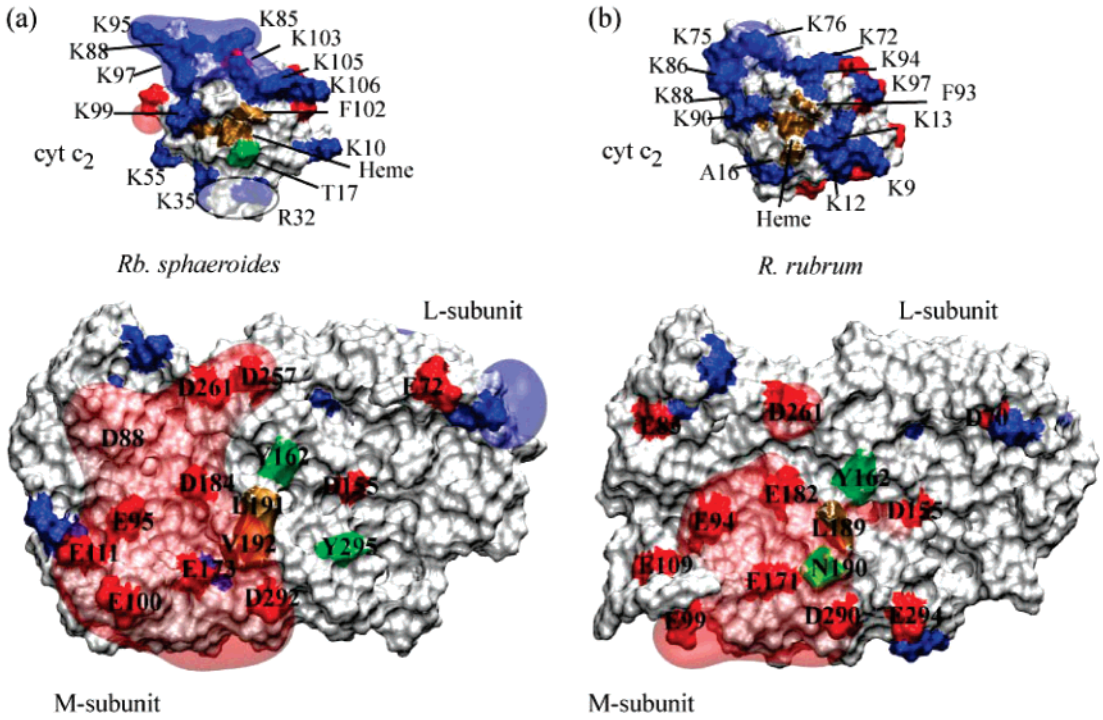

Figure 11. Interacting surfaces of the RC and cyt c_2 in (a) *Rb. sphaeroides* and (b) *R. rubrum*. The positions of the heme, direct RC−cyt c_2 contact residues, and charged docking residues of the RC and cyt c_2 are labeled. Blue indicates positively charged, red negatively charged, green polar, and brown nonpolar key docking residues. Electrostatic potentials at the docking surfaces are shown in transparent red and blue. Red corresponds to an isovalue of $-1.25k_BT/e$ and blue to an isovalue of $+1.25\ k_BT/e$. ArgC32 contained in a nonconserved loop (gray shading) in *Rb. sphaeroides* can establish a stable cation$-\pi$ interaction with TyrM295.

TABLE 1: Free Energy of the RC−Cytochrome c_2 Complex Binding Using MM-PBSA and DelPhi[a]

energy	RC−cyt c_2^{ox}	RC−cyt c_2^{red}	RC−cyt c_2^{ox}	RC−cyt c_2^{red}
$\epsilon_{internal}$	1	1	4	4
ΔE_{Coul}	-398.90 ± 5.78	-270.11 ± 6.68	-116.35 ± 2.52[b]	-82.41 ± 3.24[b]
ΔE_{VdW}	-68.96 ± 0.54	-67.56 ± 0.53		
ΔG_{solv}^{polar}	469.63 ± 5.83[b]	333.19 ± 6.19[b]	122.33 ± 2.8[b]	91.11 ± 3.03[b]
$\Delta G_{solv}^{nonpolar}$	-10.17 ± 0.06	-10.28 ± 0.06	-10.17 ± 0.06	-10.28 ± 0.06
ΔG_{eff}	-8.40 ± 1.43	-14.76 ± 1.12	-4.14 (6.03 ± 1.40[b])	-1.58 (8.70 ± 1.44[b])
$-T\Delta S_{total}$	13.9	13.9		
ΔG_{bind}	5.5	−0.9	-4.14 (6.03 ± 1.40[b])	-1.58 (8.70 ± 1.44[b])
ΔG_{bind}^{exper}		−9		−9

[a] All reported energies except contributions from entropy are averages from a 4 ns molecular dynamics trajectory for each redox state. Experimental value of free energy of binding is $\Delta G_{bind}^{exper} = -RT \ln(1/K_D) = -9$ kcal/mol.[23] [b] These values were calculated with DelPhi.

and can be attributed to changes in both the protein conformations and the charge distribution of the heme and its ligands with the oxidation state of cyt c_2. Using the crystal structure of the RC−cyt c_2 complex (PDB ID 1L9B) with both sets of charges gives much smaller but similar trends to those observed with the full ensemble of structures. By substituting the oxidized partial charges for the heme into the structures of the reduced system and comparing the energetic differences to both the reduced and the oxidized states, we find that most of the contribution arises from the structural changes, especially in the loop region containing the methionine ligand. This hypothesis is further supported by analysis of the MD equilibration runs of the RC−cyt c_2 complex. The average rmsd of the loop between helix 4 and the C-terminal helix (residues 96−104) is 2.1 Å in the reduced state compared to 2.6 Å for the complex with oxidized cyt c_2. Also, throughout the entire simulations the reduced complex has a larger number of water molecules at the interface with high residency times (data not shown). Finally, recent hydrogen exchange experiments by Englander and co-workers[88,89] on equine cyt c demonstrated that the binding affinity of the methionine ligand for the heme iron in reduced cyt c is increased by 3.2 kcal/mol, which stabilizes the loop, possibly trapping additional water molecules at the interface. Cohen and Pielak have shown that upon reduction of the heme iron the thermal stability of yeast cyt c is increased by 6.6 kcal/mol.[90]

The general redox-dependent trends in the Coulombic and polar solvation contributions in Table 1 are similar between the MM-PBSA and the DelPhi free energy estimates except that DelPhi with the PARSE charges predicts the complex with cyt c_2^{ox} to be more stable than that with cyt c_2^{red}. From a similar study of the factors contributing to the polar solvation term, we attribute this to the differences in the charge distributions between the CHARMM27 and PARSE parameters. It has been shown that the distribution of the charges and the environment surrounding the heme play an important role in the electrostatic contributions to the energy.[91] Pure electrostatic calculations (values in parenthesis in Table 1) reflect the expected slightly

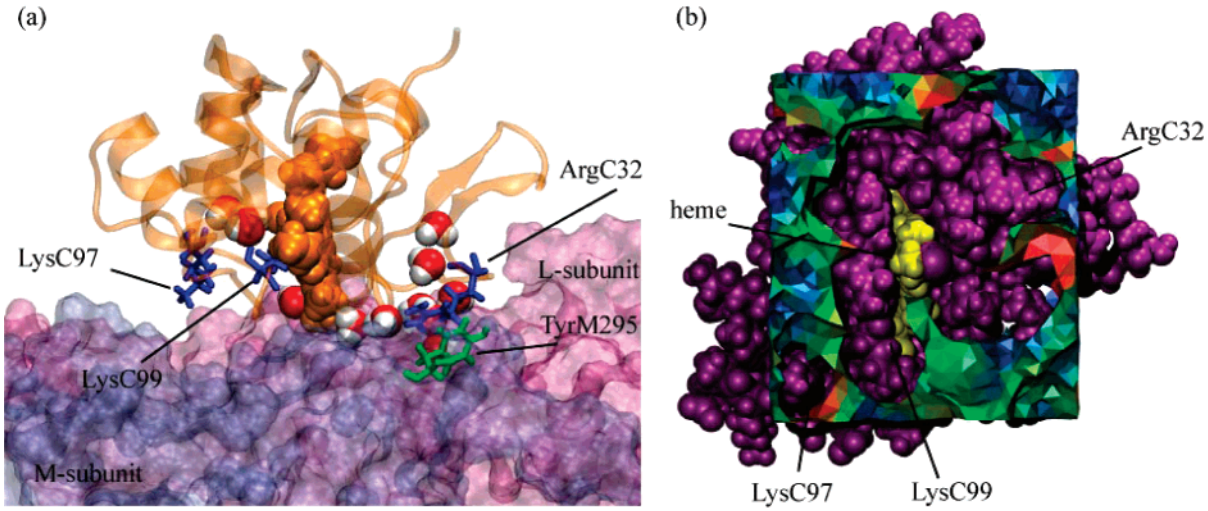

Figure 12. Solvation of the RC−cytc_2^{ox} interface. (a) Water molecules resident in the interface. (b) Bottom view of cyt c_2 showing a tessellation of the docking interface. Blue areas of the tessellation show regions of the interface that are solvated in the docked position. Green and red areas show regions that are solvated when cyt c_2 is elevated from the docked position to 3 Å.

repulsive interactions discussed earlier. Interestingly, when the nonpolar solvation component is added to the electrostatic results from DelPhi, the binding energy for the complex with cyt c_2^{red} is fairly close to the experimental measurement.

As some translational and rotational degrees of freedom are lost upon complex formation, unfavorable entropic contributions are expected. Following the approach of Finkelstein and Janin, we estimated the overall translational entropy to be $T\Delta S_{trans} \approx -TRn_{trans} \ln(\delta x/V_0) \approx -6.6$ kcal/mol, where $R = 2$ cal/(K/mol), $T = 298$ K, and $n_{trans} = 3$. The values of the rms amplitudes of translational movement in the crystal $\delta x = 0.3$ Å were taken from Axelrod and co-workers' crystallographic paper.[7] The volume available to each particle in the standard state is $V_0 = 1660$ Å^3.[63] The translational contribution is considerably higher than the value based on simple entropy of mixing. The rotational entropy was calculated as $T\Delta S_{rot} \approx -TRn_{rot} \ln(\delta\alpha/4.3) = -6.7$ kcal/mol, assuming two rotation degrees of freedom are gained upon dissociation. The rmsd of the oscillatory movements was approximated as $\delta\alpha \approx 2\delta x/d = 0.015$ radian for $\delta x = 0.3$ Å and the interface dimension, $d = 40$ Å, which was taken to be the diameter of the circle with the area of the interface of the RC−cyt c_2 complex, 1280 Å.

The vibrational contribution to entropy was estimated using the harmonic approximation. Normal modes of cyt c_2 structures in the docked and elevated conformations were calculated using the elNémo webserver.[92] In each case, only five normal modes had wavenumbers less than 208 cm$^{-1} = kT/hc$ at $T = 300$ K. The differences produced by the single wavenumber between docked and undocked structures were small. It resulted in the value of 0.1205 kcal/(mol/normal mode), which was multiplied by 5, the number of normal modes with small wavenumbers. Thus the vibrational entropy contribution was found to be small: $T\Delta S_{vib} = -0.6$ kcal/mol. Thus the total approximate value of the entropy contribution to the free energy of binding is $-T\Delta S = 13.9$ kcal/mol, which is a certain overestimation due to the presence of the membrane where the complex resides. Here we assume that the entropic contributions are independent of the redox state of cyt c_2.

The resulting value of the binding free energy of the RC−cyt c_2^{ox} was found to be $\Delta G_{bind}^{calc}(ox) \approx 5.5$ kcal/mol and of the RC−cyt c_2^{red} to be $\Delta G_{bind}^{calc}(red) \approx -0.9$ kcal/mol. The experimental value $\Delta G_{bind}^{exper} = -RT \ln(1/K_D) = -9$ kcal/mol, based on the experimentally measured dissociation constant of cyt c_2^{red} of $K_D = 0.3$ μM,[23] is clearly more negative than the values obtained from either method once the entropic contributions are added. The 1980 data of Feher and colleagues[93] indicated a 1.5-fold increase in the value of the dissociation constant for the complex with oxidized cyt c_2, which implies only a small difference in the binding energy. While the membrane and ions were included in the simulations analyzed for the MM-PBSA and DelPhi calculations, their effects on the energetic contributions to the free energy were not directly included. Neglecting the membrane probably contributes to the less favorable estimates of the absolute values of the binding free energies.

Finally, in a recent study of binding of cyt c_2 (large class) of *Rb. capsulatus* to the RC of *Rb. sphaeroides*, Cusanovich and colleagues showed, using plasmon-waveguide resonance spectroscopy, that the complex with cyt c_2^{ox} has a larger dissociation constant than that with cyt c_2^{red}.[29] While the trends observed in the MM-PBSA calculations are in agreement with the available experimental values, a more reliable comparison will require an improvement in the estimates of the entropy and membrane contributions.

3.4. Dynamic Dissociation Interface: Water Exchange. The detailed nature of the MD simulations also allows us to compare the behavior of water[94] at the complex interface in both redox states of cyt c_2. During the simulations oppositely charged residues at the interface that form transient salt bridges are partially solvated ensuring that no tight contacts are made that would hinder dissociation of cyt c_2. In the docked position of cyt c_2^{ox}, the only water molecules observed to have long residency times are in the neighborhood of the cation−π interaction, as shown in Figure 12a. In the case of cyt c_2^{red} a larger number of waters are resident especially near the loop region between residues 94 and 106 (data not shown). Tessellation[95] of the void space between the RC and cyt c_2 reveals an increase in solvation of the interface upon elevating the cyt c_2^{ox} to 3 Å above the RC (Figure 12b). Interestingly, only small regions of the cyt c_2 docking surface are solvated in the docked position (shown in blue). After elevation to 3 Å influx of water (shown in green and red) is observed in the regions surrounding a number of residues involved in important docking interactions. The behavior of the interface waters is consistent with the observed increase in motion of cyt c_2^{ox} in the docked complex to which we attribute its greater tendency to dissociate.

4. Conclusions

Using an all-atom large-scale SMD simulation in combination with evolutionary studies and binding free energy analysis we propose and explore dissociation pathways of cyt c_2 from the RC of *R. sphaeroides* embedded in a membrane without the light-harvesting complex 1. We have examined an exit pathway that could occur at the lowest elevation allowed by the surface architecture of the RC and identified conserved residues critical for this pathway. It has been found that it is energetically feasible for cyt c_2 to leave horizontally across the RC surface along the L subunit. We propose that undocking is initiated by the disruption of the conserved electrostatic contacts on the L subunit by water molecules. Van der Waals and cation$-\pi$ interactions are maintained longer during the dissociation, and the final release of cyt c_2 can be attributed to their loss. Using a model of the cyt c_2 and cyt c_1 complex for *Rb. capsulatus*, together with evolutionary analysis, we identify patterns of conservation important for the docking of cytochrome to both the RC and the cytochrome bc_1 complex. On the basis of this bioinformatic study we also predict changes in the dissociation constants for complexes of cyt c_2 of *R. rubrum* to the RC and horse cyt c to the RC, which agree with experiment or/and can be tested.

Finally, the MM-PBSA method allows us to calculate the redox-dependent binding free energies of the RC$-$cyt c_2 complex with the RC in a neutral state. According to our estimates obtained using an ensemble of structures sampled in the trajectories, the complex with oxidized cyt c_2 dissociates more readily. We attribute the reported 6 kcal/mol difference in the free energies to the reduction in the stability of cyt c_2^{ox} itself and that of the complex, which is supported from our observed redox-dependent behavior of interfacial water molecules and fluctuations in the geometry of the complex. While many of the trends in the energies can be observed using a single structure, analysis of the ensemble of structures obtained from the MD simulations makes the redox-dependent behavior more pronounced. The importance of ensemble statistics when comparing to experimental results was recently shown.[79] As the accuracy of the computational estimates of binding energy progress and high-resolution structures of other cyt c_2 complexes become available, it will be interesting to perform a comparative analysis of cyt c_2 docking to its other partners in the respiratory pathway.

Acknowledgment. The authors thank members of Professor Luthey-Schulten's group for numerous insightful discussions. Many thanks are also given to John Eargle and Dan Wright, who together with one of us (E.R.), implemented the bioinformatic tools in VMD. John Eargle also prepared Figure 12b. The authors thank Dr. Junjun Mao and Dr. Yifen Song of the Gunner Lab for kindly providing the PARSE parameters for the heme group. We are grateful to Professor Colin Wraight for stimulating discussions. This work was funded by the National Science Foundation (Grant Nos. MCB 02-35144 and MCB 04-46227). The authors also acknowledge computational time provided by the National Resource Allocation Committee (Grant No. MCA03S027P).

Supporting Information Available: A larger set of sequence-based alignments of the L and M subunits of the RC and of cyt c_1, the coordinate file of the proposed model of the cyt c_2-cyt c_1 complex for *Rb. capsulatus*, a summary of experimental results, and movies of the elevated horizontal SMD pull of cyt c_2 over the RC and of the vertical pull to 10 Å away from the RC. This material is available free of charge via the Internet at http://pubs.acs.org

References and Notes

(1) Hu, X.; Ritz, T.; Damjanović A.; Autenrieth, F.; Schulten, K. *Q. Rev. Biophys.* **2002**, *35*, 1−62.
(2) Axelrod, H. L.; Okamura, M. Y. *Photosynth. Res.* **2005**, *85*, 101−114.
(3) Berg, J. M.; Tymoczko, J.; Stryer, L. *Biochemistry*, 5th ed.; W. H. Freeman: New York, 2002.
(4) Saraste, M. *Science* **1999**, *283*, 1488−1492.
(5) *Respiration in Archaea and Bacteria*; Zannoni, D., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2004.
(6) Deisenhofer, J.; Epp, J. D. O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318*, 618−624.
(7) Axelrod, H. L.; Abresch, E. C.; Okamura, M. Y.; Yeh, A. P.; Rees, D. C.; Feher, G. *J. Mol. Biol.* **2002**, *319*, 501−515.
(8) Lange, C.; Hunte, C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2800−2805.
(9) Berry, E. A.; Huang, L. S.; Saechao, L. K.; Pon, N. G.; Valkova-Valchanova, M.; Daldal, F. *Photosynth. Res.* **2004**, *81*, 251−275.
(10) Pelletier, H.; Kraut, J. *Science* **1992**, *258*, 1748−1755.
(11) Guo, M.; Bhaskar, B.; Li, H.; Barrows, T. P.; Poulos, T. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5940−5945.
(12) Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yoshikawa, S. *Science* **1996**, *272*, 1136−1144.
(13) Yoshikawa, S.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yamashita, E.; Inoue, N.; Yao, M.; Fei, M. J.; Libeu, C. P.; Mizushima, T.; Yamaguchi, H.; Tomizaki, T.; Tsukihara, T. *Science* **1998**, *280*, 1723−1729.
(14) Roberts, V. A.; Pique, M. E. *J. Biol. Chem.* **1999**, *274*, 38051−38060.
(15) Iwata, S.; Ostermeier, C.; Ludwig, B.; Michel, H. *Nature* **1995**, *376*, 660−669.
(16) Flöck, D.; Helms, V. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 75−85.
(17) Dickerson, R. E. *Nature* **1980**, *283*, 210−212.
(18) Woese, C. R.; Gibson, J.; Fox, G. E. *Nature* **1980**, *283*, 212−214.
(19) Crowley, P. B.; Carrondo, A. M. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 603−612.
(20) Autenrieth, F.; Tajkhorshid, E.; Schulten, K.; Luthey-Schulten, Z. *J. Phys. Chem. B* **2004**, *108*, 20376−20387.
(21) Paddock, M. L.; Weber, K. H.; Chang, C.; Okamura, M. Y. *Biochemistry* **2005**, *42*, 9619−9625.
(22) Gong, X.; Paddock, M. L.; Okamura, M. Y. *Biochemistry* **2003**, *42*, 14492−14500.
(23) Tetreault, M.; Rongey, S. H.; Feher, G.; Okamura, M. Y. *Biochemistry* **2001**, *40*, 8452−8462.
(24) Farchaus, J. W.; Wachtveitl, J.; Mathis, P.; Oesterhelt, D. *Biochemistry* **1993**, *32*, 10885−10893.
(25) Wachtveitl, J.; Farchaus, J. W.; Mathis, P.; Oesterhelt, D. *Biochemistry* **1993**, *32*, 10894−10904.
(26) Tiede, D.; Dutton, P. L. In *The Photosynthetic Reaction Center*; Desenhofer, J., Norris, J. R., Eds.; Academic Press: San Diego, 1993; Vol. 1, pp 258−288.
(27) Gerencser, L.; Laczko, G.; Maroti, P. *Biochemistry* **1999**, *38*, 16866−16875.
(28) Miyashita, O.; Okamura, M. Y.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3558−3563.
(29) Devanathan, S.; Salamon, Z.; Tollin, G.; Fitch, J.; Meyer, T. E.; Cusanovich, M. A. *Biochemistry* **2004**, *43*, 16405−16415.
(30) Miyashita, O.; Onuchic, J. N.; Okamura, M. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16174−16179.
(31) Tetreault, M.; Cusanovich, M.; Meyer, T.; Axelrod, H.; Okamura, M. Y. *Biochemistry* **2002**, *41*, 5807−5815.
(32) Margoliash, E. *Proc. Natl. Acad. Sci. U.S.A.* **1963**, *50*, 672−679.
(33) Dickerson, R. E. *Sci. Am.* **1980**, *242*, 137−153.
(34) Bertini, I.; Cavallaro, G.; Rosato, A. *Chem. Rev.* **2006**, *106*, 90−115.
(35) Woese, C. R. *Microbiol. Rev.* **1987**, *51*, 221−271.
(36) Chothia, C.; Lesk, A. M. *EMBO J.* **1986**, *5*, 823−826.
(37) O'Donoghue, P.; Luthey-Schulten, Z. *J. Mol. Biol.* **2005**, *346*, 875−894.
(38) Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C. *J. Mol. Biol.* **1995**, *247*, 536−540.
(39) Sethi, A.; O'Donoghue, P.; Luthey-Schulten, Z. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4045−4050.
(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(41) Autenrieth, F.; Tajkhorshid, E.; Baudry, J.; Luthey-Schulten, Z. *J. Comput. Chem.* **2004**, *25*, 1613−1622.

(42) MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L., III.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. In *The Encyclopedia of Computational Chemistry*; von Ragué Schleyer, P., Ed.; John Wiley: New York, 1998; Vol. 1, pp 271−277.

(43) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283−312.

(44) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(45) Autenrieth, F. The photosynthetic unit of *Rhodospirillum rubrum*: A computational approach. Master's Thesis, University of Stuttgart, 2002.

(46) Humprey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.

(47) Grubmüller, H. *Solvate Version 1.0 Technical Report*; Theoretical Biophysics Group, Institute for Medical Optics, Ludwig-Maximilians University: Munich, 1996.

(48) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(49) Bayas, M. V.; Schulten, K.; Leckband, D. *Biophys. J.* **2003**, *84*, 2223−2233.

(50) Rocchia, W.; Alexov, E.; Honig, B. *J. Phys. Chem. B* **2001**, *105*, 6507−6514.

(51) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J. Comput. Chem.* **2002**, *23*, 128−137.

(52) MacKerell, A. D., Jr.; Bashford, D.; Bellot, R. L.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fisher, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(53) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.

(54) Gunner, M. R.; Honig, B. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9151−9155.

(55) Aksimentiev, A.; Schulten, K. *Biophys. J.* **2005**, *88*, 3745−3761.

(56) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III. *Acc. Chem. Res.* **2000**, *33*, 889−897.

(57) Gohlke, H.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 238−250.

(58) Zoete, V.; Meuwly, M.; Karplus, M. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 79−93.

(59) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037−10041.

(60) Minh, D. D. L.; Bui, J. M.; Chang, C.; Jain, T.; Swanson, J. M. J.; McCammon, J. A. *Biophys. J.* **2005**, *89*, L25−L27.

(61) Grünberg, R.; Nilges, M.; Leckner, J. *Structure* **2006**, *14*, 683−693.

(62) Karplus, M.; Janin, J. *Protein Eng.* **1999**, *12*, 185−186.

(63) Finkelstein, A. V.; Janin, J. *Protein Eng.* **1989**, *3*, 1−3.

(64) McQuarrie, D. A. *Statistical Mechanics*; Harper Collins Publishers: New York, 1976.

(65) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(66) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403−410.

(67) Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M.; Estreicher, A.; Gasteiger, E.; Martin, M. J.; Michoud, K.; O'Donovan, C.; Phan, I.; Pilbout, S.; Schneider, M. *Nucleic Acids Res.* **2003**, *31*, 365−370.

(68) Russell, R. B.; Barton, G. J. *Proteins: Struct., Funct., Genet.* **1992**, *14*, 309−323.

(69) Eargle, J.; Wraight, D.; Luthey-Schulten, Z. *Bioinformatics* **2006**, *22*, 504−506.

(70) Axelrod, H. L.; Feher, G.; Allen, J. P.; Chirino, A. J.; Day, M. W.; Hsu, B. T.; Rees, D. C. *Acta Crystallogr., Sect. D* **1994**, *50*, 596−602.

(71) Thompson, J. D.; Higgins, H. G.; Gibson, T. *Nucleic Acids Res.* **1994**, *22*, 4673−4680.

(72) O'Donoghue, P.; Luthey-Schulten, Z. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 550−573.

(73) Felsenstein, J. *Cladistics* **1989**, *5*, 164−166.

(74) Sokal, R. R.; Michener, C. D. *Univ. Kans. Sci. Bull.* **1958**, *28*, 1409−1438.

(75) Page, R. D. M. *Comput. Appl. Biosci.* **1996**, *12*, 357−358.

(76) Miyashita, O.; Onuchic, J. N.; Okamura, M. Y. *Biochemistry* **2003**, *42*, 11651−11660.

(77) Lin, J.; Baker, N. A.; McCammon, J. A. *Biophys. J.* **2002**, *83*, 1374−1379.

(78) Zhou, F.; Schulten, K. *J. Phys. Chem.* **1995**, *99*, 2194−2207.

(79) Suydam, I. T.; Snow, C. D.; Pande, V. S.; Boxer S. G. *Science* **2006**, *313*, 200−204.

(80) Gan, H. H.; Perlow, R. A.; Roy, S.; Ko, J.; Wu, M.; Huang, J.; Yan, S.; Nicoletta, A.; Vafai, J.; Sun, D.; Wang, L.; Noah, J. E.; Schlick, S. T. *Biophys. J.* **2002**, *83*, 2781−2791.

(81) Woese, C. R. *Syst. Appl. Microbiol.* **1984**, *5*, 315−326.

(82) Lancaster, C. R. D.; Michel, H. *Structure* **1997**, *5*, 1339−1359.

(83) Ptitsin, O. B. *J. Mol. Biol.* **1998**, *278*, 655−666.

(84) Crowley, P. B.; Golovin, A. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 231−239.

(85) Bhatia, G. E. Refinement of the crystal structure of oxidized *Rhodospirillum rubrum* cytochrome c_2. Ph.D. Thesis, University of California, San Diego, 1982.

(86) van der Wal, H. N.; Gorter, P. Y.; van Grondelle, R. *Photosynth. Res.* **1986**, *9*, 159−166.

(87) Larson, J. W.; Wraight, C. A. *Biochemistry* **2000**, *39*, 14822−14830.

(88) Xu, Y.; Mayne, L.; Englander, S. W. *Nat. Struct. Biol.* **1998**, *5*, 774−778.

(89) Krishna, M. M. G.; Maity, H.; Rumbley, J. N.; Lin, Y.; Englander, S. W. *J. Mol. Biol.* **2006**, *359*, 1410−1419.

(90) Cohen, D. S.; Pielak, G. J. *J. Am. Chem. Soc.* **1995**, *117*, 1675−1677.

(91) Mao, J.; Hauser, K.; Gunner, M. R. *Biochemistry* **2003**, *42*, 9829−9840.

(92) Suhre, K.; Sanejouand, Y. H. *Nucleic Acids Res.* **2004**, *32*, W610−W614.

(93) Rosen, D.; Okamura, M. Y.; Feher, G. *Biochemistry* **1980**, *19*, 5687−5692.

(94) Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389−415.

(95) Eargle, J.; Luthey-Schulten, Z. *Comput. Biol. Chem.* **2006**, *30*, 219−226.

(96) Roberts, E.; O'Donoghue, P.; Luthey-Schulten, Z., to be submitted for publication.

(97) Roberts, E.; Eargle, J.; Wright, D.; Luthey-Schulten, Z. *BMC Bioinf.* **2006**, *7*, 382.