Supporting Information

Pisliakov *et al.* 10.1073/pnas.0800580105

SI Text

Methods

The methods used in the present work have been described before and thus, we will present here a short summary referring the readers to the relevant detailed work in each case.

EVB and Modified Marcus Models. In studies of PTR between donors and acceptors, we describe the EVB quantum system in terms of diabatic states,

$$\begin{aligned}\Psi_i &= \mathrm{B}_1\mathrm{B}_2\ldots\mathrm{B}_i\mathrm{H}^+\ldots\mathrm{B}_n\\ \Psi_j &= \mathrm{B}_1B_2\ldots\mathrm{B}_j\mathrm{H}^+\ldots\mathrm{B}_n,\end{aligned} \qquad [5]$$

where $\mathrm{B}_i\mathrm{H}^+$ is the protonated form of the B_i protonation site (e.g., an $\mathrm{H_3O^+}$ or a protonated acid). In many cases (see, e.g., refs. 1 and 2) it has been found that a two-state model allows one to capture the physics quite well is in situations with high activation barriers (which is the case in CcO). Thus, for example, we can describe PT between two water molecules by using a two-state model:

$$\begin{aligned}\Psi_1 &= \mathrm{H_3O^+_{(a)}H_2O_{(b)}}\\ \Psi_2 &= \mathrm{H_2O_{(a)}H_3O^+_{(b)}}.\end{aligned} \qquad [6]$$

Now, the *i*th diagonal element of the Hamiltonian of this system is described by a force-field-like function that describes the bonding within donors, the bonding of the proton to the *i*th base, as well as the nonbonded interactions in the system and its interactions with the surroundings (protein or water) [the details of this force field is given elsewhere (3)]. The off-diagonal elements of the Hamiltonian, H_{12}, are described by an empirical function that is fitted to experimental information and *ab initio* calculations. The ground-state energy, E_g, is obtained by diagonalizing the EVB Hamiltonian, which, for the two-state case, gives

$$E_\mathrm{g} = \frac{1}{2}\left[(\varepsilon_1 + \varepsilon_2) - \sqrt{(\varepsilon_1 - \varepsilon_2)^2 + 4\mathrm{H}_{12}^2}\right], \qquad [7]$$

where ε_1 and ε_2 are the diabatic energies of state 1 and 2, respectively. These potential functions are given by

$$\varepsilon_i = \varepsilon_{\mathrm{SS}}^{(i)} + \varepsilon_{\mathrm{Ss}}^{(i)} + \varepsilon_{\mathrm{ss}}^{(i)} + \Delta^{(0)}. \qquad [8]$$

Here, S and s designate solute and solvent, respectively, and $\Delta^{(0)}$ is a constant, which is referred to as the gas-phase shift. In other cases, we can consider more that two EVB states. At any rate, the free energy surface associated with the above E_g is obtained by the free energy perturbation (FEP) umbrella sampling (US) procedure, described in detail elsewhere (e.g., ref. 4). This method applies a mapping potential that moves the system adiabatically from state 1 to state 2 using

$$\varepsilon_\mathrm{m} = (1 - \lambda_\mathrm{m})\varepsilon_1 + \lambda_\mathrm{m}\varepsilon_2. \qquad [9]$$

The free energy, ΔG_m, associated with changing λ_m from 0 to 1 is evaluated by the FEP procedure (4). The free energy functional that corresponds to the adiabatic ground-state surface, E_g (Eq. **5**), is then obtained by the US method as

$$\Delta g(x') = \Delta G_\mathrm{m} - \beta^{-1}\ln\langle\delta(x - x')\exp[-\beta(E_\mathrm{g}(x) - \varepsilon_\mathrm{m}(x))]\rangle_\mathrm{m}, \qquad [10]$$

where x' can be any type of coordinate to map the free energy on x. In general, the reaction coordinate x' is defined as the energy gap, $\Delta\varepsilon = \varepsilon_1 - \varepsilon_2$, where ε_1 and ε_2 are diabatic energies of the state 1 and 2. This generalized coordinate was used as the mapping coordinate in the present study.

Although the exact free energy profile for the PT process should be evaluated by the EVB or alternative approaches, it has been established both by our studies (e.g., refs. 1 and 2) and more recently by others (5) that the most important factor in determining the overall barrier for PT processes is the energy of a PT step, which is determined by the $\mathrm{pK_a}$ differences between the donor and acceptor (4).

This feature is extremely important in the cases where the microscopic calculations face major convergence problems. In such cases, we find it very useful to exploit the fact that the EVB states include the gas-phase shift of Eq. **8**, and shift the free energy surfaces that correspond to ε_i and ε_j [the microscopic Marcus' parabolas (see ref. 4)], until the ΔG_{ij} obtained by the EVB calculation is equal to the PDLD/S-LRA semimicroscopic estimate. This approach is used in the present work. This treatment yields consistent linear free energy relationships (LFER) for the PT steps, where the activation barriers are correlated with the $\mathrm{pK_a}$ difference between the donor and the acceptor.

FEP and Adiabatic Charging. Because the most important factor in the determination of the overall barrier for PT processes is the energy of charging the donor and acceptor in their given environments, it is important to have effective ways of calculating these energies. The charging free energies can be evaluated microscopically by using a standard adiabatic charging (AC) free energy perturbation (FEP) procedure (e.g., ref. 6), changing the solute residual charges from zero to their values in the solution, $\mathbf{Q}_\mathrm{S}^\mathrm{s}$, in $n + 1$ incremental steps using

$$\mathbf{Q}_\mathrm{m} = \mathbf{Q}_\mathrm{S}{}^\mathrm{S}\lambda_\mathrm{m} \qquad [11]$$

and

$$U_\mathrm{m} = U(\mathbf{Q}_\mathrm{m})_\mathrm{Ss} + U_\mathrm{ss}, \qquad [12]$$

where U_Ss and U_ss are the solute–solvent and solvent–solvent potential surfaces. The total solvation free energy is then evaluated by a FEP approach that involves gradually changing λ_m from 0 to 1. Alternatively, one could also evaluate the charging energy by using the EVB approach changing the charges of the two sates, while leaving the bonding intact. The FEP approach can also be used to evaluate the free energy of water insertion, changing water molecules to dummy atoms and this approach will be applied here.

PDLD/S-LRA Calculations. Despite the formal rigor of the microscopic methods, it is frequently found that such methods are subjected to major convergence problems, particularly when one deals with electrostatic effects in proteins interiors, and that semimacroscopic models can give more reliable results. This is true in particular with regards to the PDLD/S-LRA method (7, 8) that provides a direct link between the microscopic and macroscopic concepts. This method evaluates the change in

solvation free energies upon transfer of a given group or groups to the protein by using the effective potential (7)

$$\Delta U_{\mathrm{sol},i}^{\mathrm{w}\to\mathrm{p}} = [-\Delta G_{\mathrm{sol},i}^{\mathrm{w}} + \Delta G_{\mathrm{sol,p}}^{\mathrm{w}}(q = q_i) - \Delta G_{\mathrm{sol,p}}^{\mathrm{w}}(q = 0)] \cdot \left(\frac{1}{\varepsilon_{\mathrm{p}}} - \frac{1}{\varepsilon_{\mathrm{w}}}\right) + \Delta U_{q\mu}\frac{1}{\varepsilon_{\mathrm{p}}}, \quad [13]$$

where $\Delta G_{\mathrm{sol},\,i}^{\mathrm{w}}$ is the free energy of solvation of the ith ionizable group in water (the self-energy in water), and $\Delta G_{\mathrm{sol},\,\mathrm{p}}^{\mathrm{w}}(q = q_i)$ and $\Delta G_{\mathrm{sol},\,\mathrm{p}}^{\mathrm{w}}(q = 0)$ are the free energies of solvation of the entire protein in water with atomic charges present on the particular group ("charged state") and with atomic charges on the group set to zero ("uncharged state"), respectively. The $\Delta G_{\mathrm{sol},\,\mathrm{p}}^{\mathrm{w}}(q = 0)$ term approximates the case where the ionizable group is not in the protein cavity. $\Delta U_{q\mu}$ is the vacuum interaction between the atomic charges on the ionizable group and the permanent dipoles of the protein (represented by atomic charges), ε_{w} is the dielectric constant of water, and ε_{p} is the dielectric constant of the protein, which is basically a semimacroscopic scaling factor that accounts for the interactions that are not considered explicitly. This factor is quite different from the actual protein dielectric constant (see refs. 9 and 10).

To capture the physics of the reorganization of the protein dipoles in the charging process, it is necessary to relax the protein structure in the relevant charged and uncharged states. This is reliably done by the LRA approach (7), where the free energy for transformation between two charged states is given by

$$\Delta G_{\mathrm{sol},i}^{\mathrm{w}\to\mathrm{p}} = \frac{1}{2}[\langle \Delta U_{\mathrm{sol},i}^{\mathrm{w}\to\mathrm{p}}\rangle_{q=q_i} + \langle \Delta U_{\mathrm{sol},i}^{\mathrm{w}\to\mathrm{p}}\rangle_{q=0}], \quad [14]$$

where $\Delta U_{\mathrm{sol},\,i}^{\mathrm{w}\to\mathrm{p}}$ is the PDLD/S effective potential of Eq. **13**, the $\langle\,\rangle_{q=q_i}$ and $\langle\,\rangle_{q=0}$ terms designate an average over protein configurations generated in the charged and uncharged state of the given group, respectively. Although this approach takes the reorganization of the environment into account explicitly, it may not fully account for some effects, such as incomplete water penetration and protein reorganization. These factors are included implicitly in ε_{p}.

The basic PDLD/S-LRA calculations are performed with all protein groups (except some key residues) in their neutral form. The effect of ionizing these groups is evaluated macroscopically by finding their ionization state in a self-consistent way (9), and then evaluating the effect of these groups using interaction terms of the form $\mathrm{Q_iQ_jw_{ij}}$, where $w_{ij} = 332/(r_{ij}\varepsilon_{\mathrm{eff}}(r_{ij}))$. Here r_{ij} is the distance between the interacting groups, and $\varepsilon_{\mathrm{eff}}$ is an effective dielectric constant whose value is determined by a distance dependent function (7, 11). The justification of this approximation is discussed in detail elsewhere (7–9).

Validation Studies of Charging Energies and Effective Dielectric Constants

The selection of the most effective approach for calculations must be based on its performance on benchmarks that represent the problem at hand rather than on the formal rigorousness of the method used. This reflects the fact that simulation methods can face enormous convergence problems that cannot be assessed by the stability of the results, but rather by comparing calculated and observed results (e.g., ref. 10). Thus, we evaluated the $\mathrm{pK_a}$s of E286 and Prd by several alternative approaches and summarized the results in Table S1.

While the FEP/AC study overestimated the observed $\mathrm{pK_a}$ of E286 even with the use of the polarizable ENZYMIX force field and the insertion of water molecules, the lowest calculated $\mathrm{pK_a}$ was 13.6. Furthermore, even the use of our special approach of artificially overcharging the ionized group to induce water penetration (see ref. 12 for details) has not overcome this problem. This indicated that obtaining full convergence of the water penetration and local unfolding is very challenging. Interestingly, while the motion of water chains observed in recent MD simulations of the $\mathrm{P^-}$ region (13, 14) has no direct role in providing gates (see below), it is part of the water rearrangement and penetration that accompanies the charging of different groups. This process cannot be reliably captured by MD simulations that do not include the actual charging process. Furthermore, even with a proper FEP mapping procedure on the nanosecond time scale, it is unlikely that the correct water penetration will be simulated (12). At any rate, in view of the slow convergence, we considered the microscopic calculations mainly as general indicators of the energetics of the PT steps. Nevertheless, the microscopic calculations appeared to provide a useful way of estimating the lowest limit for the effective dielectric constant for charge-charge interactions. The estimates of $\varepsilon_{\mathrm{eff}}$ are summarized in Table S2. This table considers the change in energy of different charge configurations upon charging the specified ionized groups. This study is similar to but less rigorous than our study of charge–charge interactions in bacterial reaction centers (15).

The calculations estimate the effective dielectric "constants" for specific interactions by using the relationship

$$\Delta G_{\mathrm{calc}} = Q_iQ_j/(r_{ij}\varepsilon_{\mathrm{eff}}) = \Delta\Delta V_{\mathrm{QQ}}/\varepsilon_{\mathrm{eff}}, \quad [15]$$

where $\Delta\Delta V_{\mathrm{QQ}}$ is the change in the gas-phase electrostatic interaction between the two groups at their positions, defined by their x-ray structure. As seen from Table S2, the lower limit for $\varepsilon_{\mathrm{eff}}$ for some crucial interactions is rather small. Furthermore, if the definition of the dielectric includes major structural changes, we can obtain lower values. That is, the present study treats the torsional deformations of E and $\mathrm{P^-}$ explicitly, and this should be reflected implicitly in models that do not consider such deformation. Thus, we expect a relatively small value for the $\varepsilon_{\mathrm{eff}}$ that will convert the $\Delta\Delta V_{\mathrm{QQ}}$ contribution in the $[\mathrm{E^-\ H_3O^+\ P^-}]$ transition state in the original x-ray "open" structure to the actual free energy of this transition state in the "closed" configuration of Fig. 4*b*.

Although the calculated values of $\varepsilon_{\mathrm{eff}}$ are relatively small, they are still larger than the values obtained by macroscopic models with $\varepsilon_{\mathrm{p}} = 4$. For example, calculations of the $\mathrm{pK_a}$ of Prd are expected to yield a very negative $\mathrm{pK_a}$ for Prd by not allowing the distance between Prd and Arg-481 to change during the charging process.

Thermodynamic Cycles

This work involves different thermodynamic cycles, some of which are presented in greater detail here.

The Energetics of the $\mathrm{E^-\ H_3O^+}$ State. The energetics of the crucial $[\mathrm{E^-\ H_3O^+}]$ state can be evaluated by the cycle shown in Fig. S1.

This cycle gives the following expression:

$$(\Delta G^{\neq})^{\mathrm{p}} \cong 1 + \Delta G_{\mathrm{w}} + (\Delta\Delta G_{\mathrm{sol}})_{\mathrm{np}\to\mathrm{p}}^{\mathrm{w}\to\mathrm{p}} + \Delta\Delta G(\mathrm{EH_{sol}})_{\mathrm{po}\to\mathrm{np}}^{\mathrm{w}\to\mathrm{p}} + \Delta G(\mathrm{w}_1)_{\mathrm{np}}^{\mathrm{w}\to\mathrm{p}} + \Delta G(\mathrm{w}_2)_{\mathrm{po}}^{\mathrm{w}\to\mathrm{p}} + \Delta G_{\mathrm{conf}}, \quad [16]$$

where "po" and "np" designate polar and nonpolar respectively, ΔG_{w} is the free energy of the given step in water, and ΔG_{conf} is the free energy of the torsional deformations.

Water Insertion. The energetics of the PT is not determined directly by the connectivity of water molecules (as could have been understood from some works (e.g., refs. 13 and 14) nor even by the hydrogen-bonding pattern of the unprotonated water molecules (see ref. 1). When a water molecule serves as a site for

the proton on the way between a donor and acceptor, the key issues are the energetics of placing this water at the specific site and, much more importantly, the energy of protonating it. These energy contributions cannot be quantified by running long MD simulations (13, 16) but by actual free energy calculations (see below). It should also be clarified that the insertion of water molecules at available cavities and the relaxation of the orientation of these water molecules by MD runs has been part of our simulation protocol since 1984. In fact, inserting simplified water molecules has been a part of our PDLD simulation protocol since 1976 (17). However, such procedures do not assess the energy of placing water molecules in hydrophobic cavities. More specifically, inserting water molecules in the hydrophobic region between E and P^- and running long MD simulations (e.g., refs. 13 and 14) is not likely to be highly informative because the water molecules will most probably be trapped in this region even in ns simulations, regardless of if they are more stable in the bulk water. Thus, even interesting identification of hydrogen-bonding formation (13) is not so informative, because it does not take into account the chance that water molecules will be in the given region and, because the hydrogen bonding does not tell us about the PT energetics (1, 10, 18). Here, it is crucial to use some energy-based insertion approach.

As should be clear from the main text, the energetics of inserting water molecules can change the balance between the barriers for different paths. Thus, it might be quite important to try to estimate the free energy of inserting water molecules at critical sites. As stated above, other studies have explored some elements of this issue (19–22), but not in the framework of a thermodynamic cycle that gives the overall actual barriers for the PT steps. In other words, although insightful studies have looked for the stability of inserted water molecules in different sites by examining whether these molecule would stay there after insertion, we are not aware of studies that evaluated the free energy of transferring the water molecules from the bulk solvent to the specific assumed sites. Similarly, instructive studies that explored the energetics of orienting neutral water molecules are not directly related to the energy of the PT process, which must consider the energy of protonated water molecules (see discussion in ref. 1). It is also useful to realize that our EVB approach automatically evaluates the free energy of orienting the water molecule in either the protonated and unprotonated states. Thus, the only missing element in our previous treatment of the role of water molecules in PT in proteins was the energy of inserting water molecules to critical sites.

Our procedure of evaluating the free energy of inserting water molecules involves the cycle described in Fig. S3. This cycle provides the free energy of inserting nonpolar or polar water molecule (both are needed in different cases).

The practical evaluation of the insertion free energy also involves the use of a specialized restraint that allows us to consistently obtain the overall free energy of the PT step. That is, $3n$ dummy atoms (representing n water molecules) are inserted between the given donor and acceptor atoms, and held by a weak distance constraint that keeps them ≈3 Å away both from each other and from the donor or the acceptor. The dummy atoms are then sequentially mutated into n water molecules (according to the cycle of Fig. 2), both in the protein and in water, and the specific PT step was simulated. The same procedure is repeated in the bulk solvent. The restraint is then released in the transition state as well as in the ground state, both in the protein and in water, leaving only a small cage constraint on each water molecule (see ref. 23 on the nature of this restraint). The overall cycle gives the free energy of inserting the water molecules and evaluating the energy of the given PT process (relative to the same PT in water).

Mutational Analysis

The present study focuses on the reaction paths in the native CcO. However, the reproduction of the key mutational effects can provide a useful check on the predicted reaction paths. Although more systematic studies are left to subsequent works, we comment here at a preliminary level on some key mutational effects.

The N139D Mutation. This mutation (24) changes the relative heights of the barrier for PT to P^- and PT to the Bn. Our previous study (25) assumed that the rate of PT to the Bn is controlled by an electron transfer (ET) from heme a to heme a_3. It was found that the interaction between the negative charge of the 139D mutant and the charge generated in the TS of the concerted path increases the activation barrier by ≈1–2 kcal/mol due to the stabilization of the state with a protonated water just before E. The present study can account for the mutational effect by the destabilization of the [E^- H_2O PH] state due to the electrostatic repulsion between D139 and E^-. This requires, however, that the above state will be a high-energy state that is trapped (because of the need for additional energy for the ET from heme a to heme a_3) for a longer time than the time needed for a PT from E to B^-. More careful examination of the nature of the effect of the mutations will be left to a stage, where we have gained more quantitative knowledge of the energy of different states.

The W172F Mutation. The W172F mutation leads to a significant change in the pumping of CcO and can be interpreted (26) as an increase in the barrier for PT between the E and P. The reproduction of the change in the observed barrier may provide an instructive validation of the present model. Unfortunately, the mutation can change the energy of water insertion (ΔG_{ins}), the energy of rotating E and P (ΔG_{conf}) and the solvation change ($\Delta\Delta G_{solv}$). Thus, we leave such detailed calculations to a subsequent study.

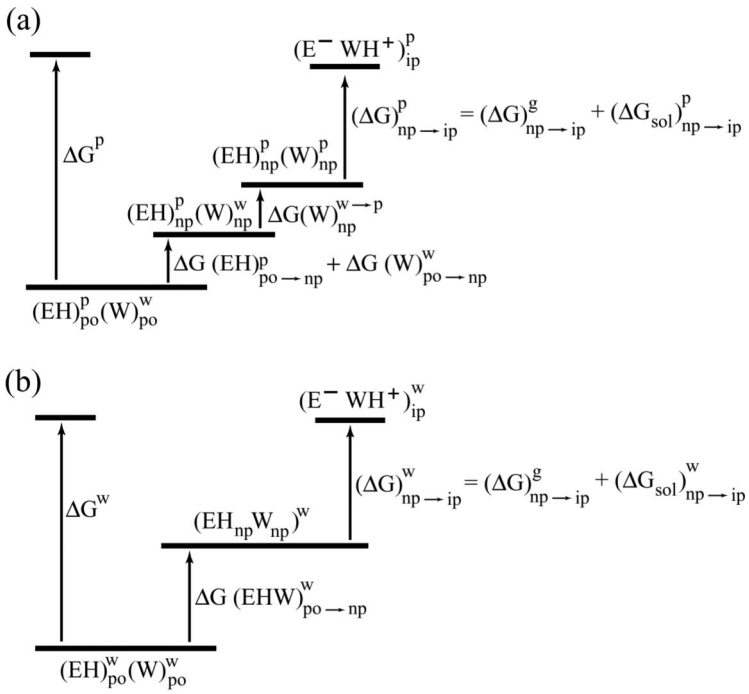
The E286A/I112E Mutation. This mutation (27) results in the shift of the primary donor further from P^-. This reduces the pumping and probably slows down the rate of the PT to P^-. The closest distance between the I112 and Prd is 12 Å. Thus, the PT barrier in this mutant is probably similar to that of the open configuration of the E286-Prd pair. Here, it would be important to evaluate the height of the barrier by careful simulations, but at present it seems that our studies overestimate the energy of the open [E^- W_1H^+ W_2P^-] configuration.

Quantum Mechanical Validations

The present study explored configurations where two acidic groups are bridged by a water molecule and where the PT process corresponds to the transfer from the [EH W P^-] to the [E^- W PH] configuration. Because this process may involve a reduced barrier by a concerted path (2), it is important to verify the EVB calculations by related hybrid *ab initio* quantum mechanical calculations, for which we used the B3LYP method (28, 29) combined with the 6–311++G** basis sets, and the PCM model (30) for the solvent calculations. The validation study explored the energetics of PT between two formic acids bridged by a water molecule. The corresponding results for the gas-phase and the solution reactions are described in Fig. S2. The main point that emerges from the calculation is the finding that, even at a distance of 5 Å between the oxygen atoms of the acids, the barrier is significant. This result agrees well with the EVB estimate of this barrier. Furthermore, the barrier in solution can easily be estimated from the pK_a difference between the donor and the protonated acceptor (e.g., the first formic acid and the protonated water molecule) and the electrostatic stabilization by

the formate ion of the second acid. Because the same results are obtained by the EVB calculations, we can conclude that we are not missing any unexpected delocalization effects by using our EVB approach.

1. Braun-Sand S, Burykin A, Chu ZT, Warshel A (2005) Realistic simulations of proton transport along the gramicidin channel: Demonstrating the importance of solvation effects. *J Phys Chem B* 109:583-592.
2. Braun-Sand S, Strajbl M, Warshel A (2004) Studies of proton translocations in biological systems: Simulating proton transport in carbonic anhydrase by EVB based models. *Biophys J* 87:2221-2239.
3. Burykin A, Warshel A (2003) What really prevents proton transport through aquaporin? Charge self-energy versus proton wire proposals. *Biophys J* 85:3696-3706.
4. Warshel A (1991) *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (Wiley, New York).
5. Riccardi D, Konig P, Guo H, Cui Q (2008) Proton transfer in carbonic anhydrase is controlled by electrostatics rather than the orientation of the acceptor. *Biochemistry* 47:2369-2378.
6. Warshel A, Sussman F, King G (1986) Free energy of charges in solvated proteins: Microscopic calculations using a reversible charging process. *Biochemistry* 25:8368-8372.
7. Lee FS, Chu ZT, Warshel A (1993) Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. *J Comput Chem* 14:161-185.
8. Sham YY, Chu ZT, Warshel A (1997) Consistent calculations of pK_a's of ionizable residues in proteins: Semi-microscopic and microscopic approaches. *J Phys Chem B* 101:4458-4472.
9. Schutz CN, Warshel A (2001) What are the dielectric 'constants' of proteins and how to validate electrostatic models. *Proteins Struct Funct Genet* 44:400-417.
10. Warshel A, Sharma PK, Kato M, Parson WW (2006) Modeling electrostatic effects in proteins. *Biochim Biophys Acta* 1764:1647-1676.
11. Warshel A, Russell ST, Churg AK (1984) Macroscopic models for studies of electrostatic interactions in proteins: Limitations and applicability. *Proc Natl Acad Sci USA* 81:4785-4789.
12. Kato M, Warshel A (2006) Using a charging coordinate in studies of ionization induced partial unfolding. *J Phys Chem B* 110:11566-11570.
13. Seibold SA, Mills DA, Ferguson-Miller S, Cukier RI (2005) Water chain formation and possible proton pumping routes in Rhodobacter sphaeroides cytochrome c oxidase: A molecular dynamics comparison of the wild type and R481K mutant. *Biochemistry* 44:10475-10485.
14. Wikstrom M, *et al.* (2005) Gating of proton and water transfer in the respiratory enzyme cytochrome c oxidase. *Proc Natl Acad Sci USA* 102:10478-10481.
15. Sham YY, Muegge I, Warshel A (1998) The effect of protein relaxation on charge-charge interactions and dielectric constants of proteins. *Biophys J* 74:1744-1753.
16. Hosler JP, Ferguson-Miller S, Mills DA (2006) Energy transduction: Proton transfer through the respiratory complexes. *Annu Rev Biochem* 75:165-187.
17. Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 103:227-249.
18. Burykin A, Warshel A (2004) On the origin of the electrostatic barrier for proton transport in aquaporin. *FEBS Lett* 570:41-46.
19. Wikstrom M, Verkhovsky MI, Hummer G (2003) Water-gated mechanism of proton translocation by cytochrome c oxidase. *Biochim Biophys Acta Bioenerg* 1604:61-65.
20. Olkhova E, Hutter MC, Lill MA, Helms V, Michel H (2004) Dynamic water networks in cytochrome c oxidase from Paracoccus denitrificans investigated by molecular dynamics simulations. *Biophys J* 86:1873-1889.
21. Xu JC, Voth GA (2005) Computer simulation of explicit proton translocation in cytochrome c oxidase: The D-pathway. *Proc Natl Acad Sci USA* 102:6795-6800.
22. Popovic DM, Stuchebrukhov AA (2004) Electrostatic study of the proton pumping mechanism in bovine heart cytochrome c oxidase. *J Am Chem Soc* 126:1858-1871.
23. Štrajbl M, Sham YY, Villà J, Chu ZT, Warshel A (2000) Calculations of activation entropies of chemical reactions in solution. *J Phys Chem B* 104:4578-4584.
24. Pawate AS, *et al.* (2002) A mutation in subunit I of cytochrome oxidase from Rhodobacter sphaeroides results in an increase in steady-state activity but completely eliminates proton pumping. *Biochemistry* 41:13417-13423.
25. Olsson MHM, Warshel A (2006) Monte Carlo simulations of proton pumps: On the working principles of the biological valve that controls proton pumping in cytochrome c oxidase. *Proc Natl Acad Sci USA* 103:6500-6505.
26. Ribacka C, *et al.* (2005) An elementary reaction step of the proton pump is revealed by mutation of tryptophan-164 to phenylalanine in cytochrome c oxidase from Paracoccus denitrificans. *Biochemistry* 44:16502-16512.
27. Gilderson G, Aagaard A, Brzezinski P (2002) Relocation of an internal proton donor in cytochrome c oxidase results in an altered pK(a) and a non-integer pumping stoichiometry. *Biophys Chem* 98:105-114.
28. Becke AD (1993) A new mixing of Hartree–Fock and local density-functional theories. *J Chem Phys* 98:1372-1377.
29. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A* 38:3098-3100.
30. Barone V, Cossi M, Tomasi J (1998) Geometry optimization of molecular structures in solution by the polarizable continuum model. *J Comput Chem* 19:404–417.



$$\Delta G^p \approx \Delta G^w + (\Delta G_{sol})^{w\rightarrow p}_{np\rightarrow ip} + \Delta G(W)^{w\rightarrow p}_{np} + \Delta\Delta G(EH)^{w\rightarrow p}_{po\rightarrow np}$$

Fig. S1. Thermodynamic cycles for the primary PT in CcO. ''po,'' ''np,'' and ''ip'' designate polar, nonpolar, and an ion-pair, respectively. The energetics in the protein and in water are described in *a* and *b*, respectively. ''ΔG_{sol}'' designates solvation free energy, and ''p'' and ''w'' designate protein and water, respectively. It is assumed that $\Delta G(\text{EH W})^w_{po\rightarrow np} \approx \Delta G(\text{EH})^w_{po\rightarrow np} + \Delta G(\text{W})^w_{po\rightarrow np}$.

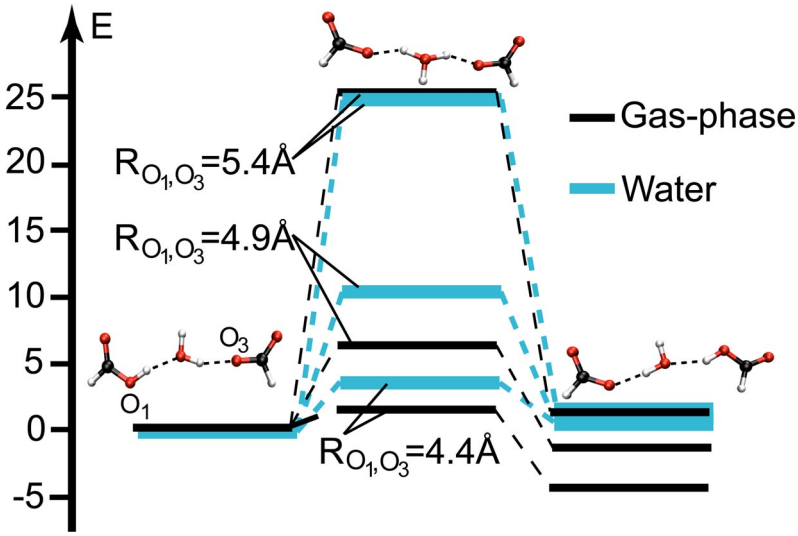
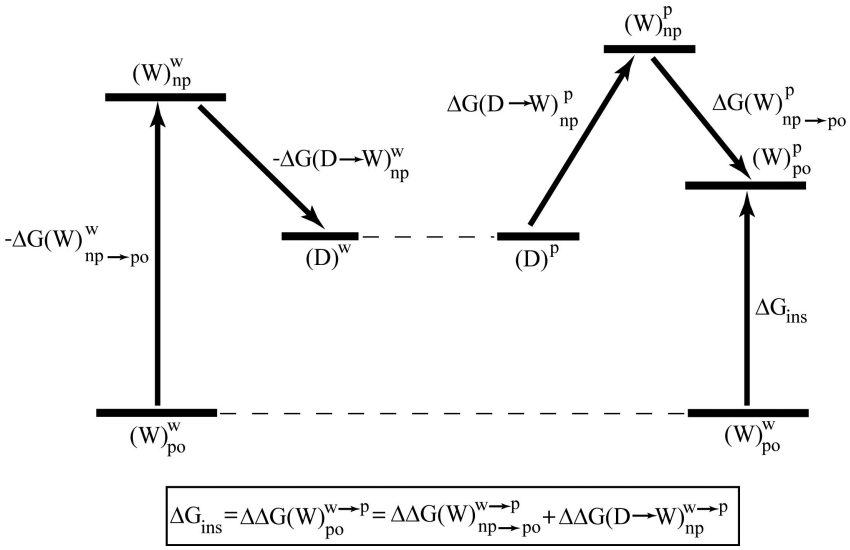


Fig. S2. Energetics of PT between two acids bridged by a water molecule.



$$\Delta G_{ins} = \Delta\Delta G(W)^{w \rightarrow p}_{po} = \Delta\Delta G(W)^{w \rightarrow p}_{np \rightarrow po} + \Delta\Delta G(D \rightarrow W)^{w \rightarrow p}_{np}$$

Fig. S3. Thermodynamic cycle for the water insertion process. ''po,'' ''np,'' and ''D'' designate polar, nonpolar, and dummy atoms, respectively. The states in water [e.g., $(W)^{w}_{po}$] also include implicitly the protein as a spectator. The cycle does not include the restraint treatment described in the text.

Table S1. The energetics of some charge configurations in CcO

System	Approach	$\Delta\Delta G_{sol}^{AH\rightarrow A^-}$	$pK_{a,\ calc}$	$pK_{a,\ obs}$
E286[†]	FEP/AC (pol)[‡]	24.4	21.7	9.4[§]
E286[¶]	FEP/AC (pol)	17.2	16.5	9.4[§]
E286[¶]	FEP/AC (pol**)	13.3	13.6	9.4[§]
E286[†]	PDLD/S-LRA	10.0	11.2	9.4[§]
E286[‡]	PDLD/S-LRA	9.0	10.5	9.4[§]
PrdH	FEP/AC (pol)	−3.0	1.8	>5[‖]
PrdH[‖]	PDLD/S-LRA	−8.6	−2.2	>5

Energies are in kcal/mol; pK_as are in pK_a units.

[†]E286 in the ″open″ configuration of Fig. 4*a*.

[‡]″pol″ designates polarizable force field that does not include the induced dipoles force, and ″pol**″ designates a polarizable force field that includes the forces associated with the induced dipoles.

[§]Taken from Namslauer and Brzezinski [Namslauer A, Brzezinski P (2004) *FEBS Lett* 567:103–110].

[¶]E286 in the ″open″ configuration of Fig. 4*a* with water molecules inserted near this acid.

[‖]The pump function requires that $pK_a(Prd) < pH_p$.

Table S2. Estimating ε_{eff} in CcO

System	Method	$\Delta\Delta V_{QQ}$	$\Delta\Delta G$	ε_{eff} †	$\varepsilon_{eff,\ obs}$
[E^- H_3O^+] (P^-)‡	FEP/AC	17.2	2.0	8.5	
[E^- H_3O^+] (P^-)‡	PDLD/S-LRA	17.2	1.4	12.2	
[E^- H_3O^+] (P^-)§	PDLD/S-LRA	14.8	1.2	12.3	
($D139^-$) [E^- H_3O^+]‡	FEP/AC	3.6	0.6	6.0	7.5¶
($D139^-$) [E^- H_3O^+]‡	PDLD/S-LRA	3.6	0.43	8.3	7.5¶
($D139^-$) [E^-]	FEP/AC	15.2	1.9	7.8	7.5¶
($D139^-$) [E^-]	PDLD/S-LRA	15.2	2.0	7.5	
[(P^-)] (Arg-481)	PDLD/S-LRA	108.1	13.0	8.3	
[H_3O^+]‡ (P^-)	PDLD/S-LRA	70.2	6.0	11.0	

Energies are in kcal/mol.

†ε_{eff} is evaluated through Eq. **15** as the effect of the charge on the group within small brackets [e.g., (P^-)] on the charging of the groups within square brackets (e.g., [E^- H_3O^+]).

‡In the closed configuration of Fig. 4*b*.

§In the open configuration of Fig. 4*a*.

¶Based on an estimated change of 1.6 pK_a units [Namslauer A, Pawatet AS, Gennis R, Brzezinski P (2003) *Proc Natl Acad Sci USA* 100:15543–15547] for a distance of ≈20 Å between D139 and E286.