PROTEINS: Structure, Function, and Bioinformatics 64:829–844 (2006)

REVIEW

The Barrier for Proton Transport in Aquaporins as a Challenge for Electrostatic Models: The Role of Protein Relaxation in Mutational Calculations

Mitsunori Kato[1], Andrei V. Pisliakov[1], and Arieh Warshel[1]*
[1]*Department of Chemistry, University of Southern California, Los Angeles, California*

ABSTRACT **The origin of the barrier for proton transport through the aquaporin channel is a problem of general interest. It is becoming increasingly clear that this barrier is not attributable to the orientation of the water molecules across the channel but rather to the electrostatic penalty for moving the proton charge to the center of the channel. However, the reason for the high electrostatic barrier is still rather controversial. It has been argued by some workers that the barrier is due to the so-called NPA motif and/or to the helix macrodipole or to other specific elements. However, our works indicated that the main reason for the high barrier is the loss of the generalized solvation upon moving the proton charge from the bulk to the center of the channel and that this does not reflect a specific repulsive electrostatic interaction but the absence of sufficient electrostatic stabilization. At this stage it seems that the elucidation and clarification of the origin of the electrostatic barrier can serve as an instructive test case for electrostatic models. Thus, we reexamine the free-energy surface for proton transport in aquaporins using the microscopic free-energy perturbation/umbrella sampling (FEP/US) and the empirical valence bond/umbrella sampling (EVB/US) methods as well as the semimacroscopic protein dipole Langevin dipole model in its linear response approximation version (the PDLD/S-LRA). These extensive studies help to clarify the nature of the barrier and to establish the "reduced solvation effect" as the primary source of this barrier. That is, it is found that the barrier is associated with the loss of the generalized solvation energy (which includes of course all electrostatic effects) upon moving the proton charge from the bulk solvent to the center of the channel. It is also demonstrated that the residues in the NPA region and the helix dipole cannot be considered as the main reasons for the electrostatic barrier. Furthermore, our microscopic and semimacroscopic studies clarify the problems with incomplete alternative calculations, illustrating that the effects of various electrostatic elements are drastically overestimated by macroscopic calculations that use a low dielectric constant and do not consider the protein reorganization. Similarly, it is pointed out that microscopic potential of mean force calculations that do not evaluate the electrostatic barrier relative to the bulk water cannot be used to establish the origin of the electrostatic barrier. The relationship between the present study and calculations of pK_as in protein interiors is clarified, pointing out that approaches that are applied to study the aquaporin barrier should be validated by pK_as calculations. Such calculations also help to clarify the crucial role of solvation energies in establishing the barrier in aquaporins. Proteins 2006;64:829–844.**

Key words: aquaporins; proton transport; membrane proteins; water channel; free-energy perturbation; EVB; PMF; LRA; electrostatic calculations

INTRODUCTION

The discovery of aquaporins and their remarkable role in conducting water molecules through cell membranes has attracted major interest in recent year (e.g., Refs. 1–3) One of the important questions that has been raised by the discovery of aquaporins is the origin for the blockage of protons by the aquaporins channel. This issue has been a subject of major current interest,[4] and is continuing to attract the attention of the biophysical community.[5–14] Early studies (e.g., Refs. 5, 9) suggested that the blockage is attributable to water orientational effects that disrupt the Grotthuss mechanism.[15–17] However, recent works[6,8,10,11,14,18] reflect the view that the blockage is

Grant sponsor: National Institutes of Health; Grant number: GM40283.

*Correspondence to: Arieh Warshel, Department of Chemistry, SGM 418, University of Southern California, 3620 McClintock Avenue, Los Angeles, CA 90089-1062. E-mail: warshel@usc.edu

Received 30 November 2005; Revised 21 February 2006; Accepted 1 March 2006

Published online 15 June 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21012

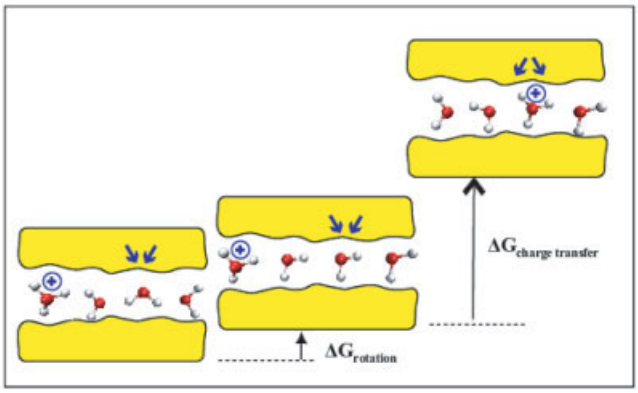

Fig. 1. Illustrating the idea that the barrier in aquaporin is not due to the orientation of the water molecules but to the electrostatic barrier for the transfer of the proton charge. The figure divides the overall process of PTR to the center of the channel into two steps; first, the orientation of the water molecules and then transferring the proton charge to the center of the channel while overcoming the electrostatic barrier. The figure ignores the fact that a complete orientation of the water molecules is not needed (see Refs. 23, 46) and illustrates schematically the finding that the orientation needed for the actual PTR does not cost significant energy relative to the energy of moving the proton charge against the electrostatic barrier. The origin of this barrier is the subject of the present work (which is not necessarily the NPA dipoles).

attributable to the electrostatic barrier, in agreement with our early general proposal,[19,20] which argued that proton transport (PTR) in proteins is controlled by the electrostatic barriers associated with the movement of the proton charge through its specific path. The current consensus is summarized in Figure 1, which illustrates schematically the idea that the energy invested in orienting the water molecules is much smaller than the energy required to move a positive charge of the proton through the channel.

Assuming that the above point is generally accepted, we can move to our main subject, which remains quite controversial, namely, the origin of the electrostatic barrier and its magnitude. This controversy also seems to reflect misunderstandings of the requirements from consistent electrostatic calculations in protein. Some workers have attributed the barrier to special structural elements[6,8] and in particular to the so-called NPA motif,[4,8,14,18] to ionized residues,[14] and/or to the helix dipoles.[10,11] However, Burykin and Warshel[12,13] (referred to here as BW) concluded that although the electrostatic barrier reflects all the electrostatic contributions of the channel (polar and nonpolar groups) the barrier will remain very high even when these contributions are removed. The different views can be summarized by the schematic drawings of Figure 2, which presents crucial modifications and clarifications (see below) of a similar illustration that was presented recently in Ref. 10.

The figure presents: a) the early view that barrier is due to the orientation constraint imposed on the water molecules by the NPA motif;[5–9] b) the misunderstanding of the BW model, where the cannel is presumably described by a low dielectric environment,[10] a view that has never been proposed by BW,[12,13] who actually emphasized that the overall effect of the channel can be represented by a relatively high effective dielectric $\epsilon = 7$ (when the dielectric is described by the effective solvation and, of course, represents the channel interior by polar dipole rather than a low dielectric); c) the same model as in b but now with the more reasonable assumption that the pore is represented by a high dielectric or by a simplified solvent model, thus presenting a somewhat less unjustified view of the BW model; d) the actual proposal of BW, who considered the complete set of all the electrostatic elements in the channel in a consistent way (see below) and examined the overall effect in substituting the solvation by the bulk water; e) the same as in model d but with fixed protein [i.e., the model does not allow the protein dipoles to rearrange during the charging process (e.g., Refs. 8, 11)]; and f) the idea that the barrier is due to helix dipoles.[10,11]

The validation of the different proposals has been perhaps slowed because of the implicit assumption that electrostatic effects cannot control PTR. However, now when EVB and other models demonstrated that the PTR profile follows the electrostatic profile (at least in aquaporin), it became possible to focus on the nature of the electrostatic effect and to try to reach a consensus on this issue. Unfortunately, the present situation may still reflect unfamiliarity with the previous progress in electrostatic calculations in proteins and with the early conceptual advances in the description of solvation of charges by proteins.

Part of the reason for having different views about the origin of the electrostatic barrier is associated with the use of different methods, including some that have not been validated in studies of known electrostatic contributions and pK_a calculations. The problem also reflects the fact that most of the reported studies have not examined the actual contributions of different residues by removing these residues. Another problem may be associated with the perception that potential of mean force (PMF) calculations should provide quantitative results about the electrostatic profile in aquaporin and about the effects of mutations.

The present article will revisit recent calculations of the electrostatic profile in aquaporin and the effect of mutations on this profile. The article will describe semimacroscopic and microscopic calculations of the effect of different factors emphasizing the reliability of the calculations and the problems with alternative approaches. In particular, it will illustrate the problems with incomplete PMF calculations of mutational effects as well as with macroscopic calculations that use a very low dielectric and keep the protein in a single geometry during the charging process. It will be clarified that such models provide unrealistic results for the energetics of charges by proteins and thus for the energetics of H_3O^+. The paper will also clarify some misunderstandings about the concept of "solvation" by proteins. It will be pointed out that "solvation by the protein" and "solvation substitution" [12,21] are some of the most fundamental and key electrostatic effects rather than special nonelectrostatic effects as may be understood from some recent works.[10,14]

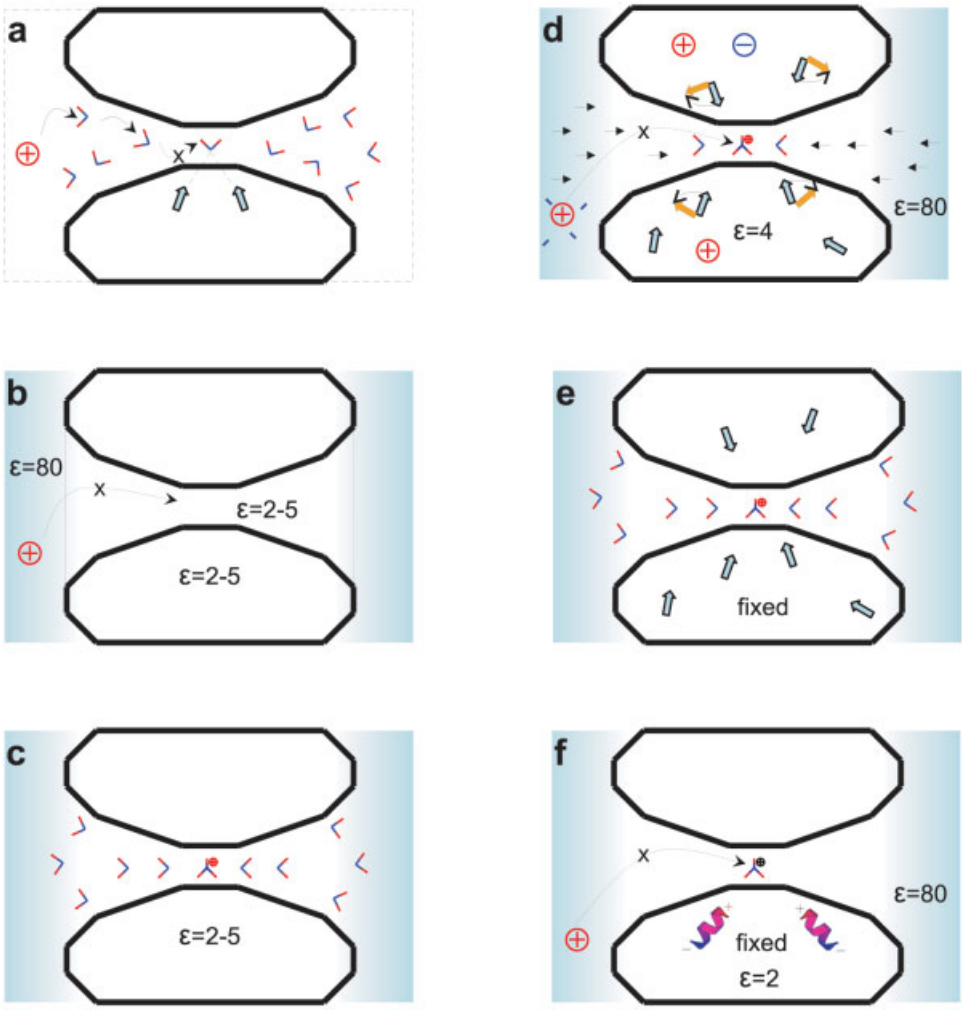

Fig. 2. A schematic description of different proposals for the origin of the barrier for PTR in aquaporin. The figure presents: **a)** the well understood view that the barrier is due to the orientation constraint imposed on the water molecules by the NPA motif;[5,9] **b)** the misunderstanding[10] of the model of Burykin and Warshel **(**BW),[12,13] where the channel is presumably described by a low dielectric environment (a view that was never proposed by BW who actually emphasized that the overall effect of the channel can be represented by a relatively high effective dielectric $\epsilon = 7$; when the dielectric is described by the effective solvation); **c)** the same model as in b but now with the more reasonable assumption that the pore is represented by a high dielectric or by a simplified solvent model, thus presenting a slightly less incorrect view of the BW model; **d)** the actual proposal of BW, who considered the complete set of all the electrostatic elements in the channel in a consistent way (see below) and examined the overall effect in substituting the solvation by the bulk water; **e)** the same as in model d but with fixed protein (i.e., the model does not allow the protein dipoles to rearrange during the charging process (e.g., Refs. 8, 11); and **f)** the idea that the barrier is due to helix dipoles.[10,11]

METHODS

In this article, we will focus on the electrostatic profile, accepting (at least as a working hypothesis) the idea[20] that the PTR profile follows the electrostatic profile when the barrier is relatively high (see Refs. 10, 12, 13, 19, 20, 22 for general consideration, and Ref. 23 for quantitative evidences). Moreover, because we are interested in proton transfer processes we should focus on the electrostatic contributions to the protonation energies (and the corresponding pK_a values) of the different protonation sites in the water chain (keeping in mind that these protonation energies are determined by the electrostatic profile).

In searching for accurate models for the evaluation of pK_a values in heterogeneous and sometimes highly charged protein environments, it is important to realize that some formally rigorous methods do not yet provide converging results. Thus, it is crucial to validate the given approach by pK_a calculations or to consider previously reported validation results, while being very critical about the benchmark used and its relevance to the current problem. With this in mind, we will consider here three approaches: the microscopic all atom PMF calculations based on free-energy perturbation (FEP) adiabatic charging (AC)[24]/umbrella sampling (US) treatment (FEP/US), the empirical valence bond (EVB)/umbrella sampling (EVB/US)

approach and the semimacroscopic protein dipole Langevin dipole model in its linear response approximation version (PDLD/S-LRA) approach.[25,26] We will also comment on the common approach of dragging the probe charge along the channel. These approaches will be briefly reviewed below.

The FEP/US Method

One of the most rigorous ways of evaluating electrostatic free energies is the FEP method.[27,28] The FEP method can be used to evaluate the free energy of moving a charge across an ion channel by combining it with a US approach.[29] This FEP/US changes the charge of the system from 0 to Q_0 in an AC process and evaluates the corresponding free energy by using a mapping potential of the form:

$$V_m(\lambda_m) = (1 - \lambda_m)V_b + \lambda_m V_a + E_{\text{cons}}(z) \quad (1)$$

where V_a and V_b are the potential of the solute–solvent interaction for $Q = 0$ and $Q = Q_0$, respectively, and where λ_m changes by n + 1 constant increments from zero to one. The additional potential $E_{\text{cons}}(z)$ keeps the system at a given region of the channel axis, z, and is given by $E_{\text{cons}}(z) = K(z - z_0)^2$. The free energy associated with the charging process is then evaluated by (e.g., Refs. 24, 30):

$$\Delta\Delta G(\lambda_m \rightarrow \lambda_{m+1}) = -\frac{1}{\beta}\ln[\langle\exp\{-(V_{m+1} - V_m)\beta\}\rangle_m], \quad (2)$$

$$\Delta G(V_a \rightarrow V_b) = \Delta G(\lambda_0 \rightarrow \lambda_n) = \sum_{m=0}^{n-1} \Delta\Delta G(\lambda_m \rightarrow \lambda_{m+1})$$

where λ_m changes in n + 1 steps between zero and one.

Combining the FEP and the US approach by the same implementation used in our studies of electron transfer reactions[29,31] and in the evaluation of free energies by the EVB approach,[32] we can obtain the free-energy probability at the Q_0 state for different points along the channel axis by using:[33,34]

$$\Delta g_b(z)_{AC} = -\frac{1}{\beta}\ln[\exp\{-\Delta G(V_a \rightarrow V_b)\beta\} \times \langle\delta(z' - z)\exp\{E_{\text{cons}}(z)\beta\}\rangle_b]. \quad (3)$$

It is useful to mention at this point that in many cases it is very hard to perform converging FEP calculations (e.g., binding of large ligands). In such cases, it is extremely useful to estimate the free energy of biological processes by an equation derived in[25] and used in studies of ligand binding to proteins (e.g., Refs. 25, 26). This equation expresses the free energy associated with changing the potential of the system from U_1 to U_2 by:

$$\Delta G(U_1 \rightarrow U_2) = \frac{1}{2}(\langle U_2 - U_1\rangle_1 + \langle U_2 - U_1\rangle_2). \quad (4)$$

The derivation of this equation was based on the assumption that the linear response approximation (LRA)[35] is valid. Namely, the protein and solvent environments respond linearly to the force associated with the given process. This assumption is, in fact, the basis of macroscopic theory where the free energy of charging a positive ion is given by the well-known result $\Delta G = \frac{1}{2}\langle U\rangle_{Q=1}$, where $\langle U\rangle_{Q=1}$ is the average of the electrostatic potential of the given charge.[36] Although it is not obvious that the LRA can provide a reliable way of describing the energetics of macromolecules or of realistic molecular systems, it was found by simulation studies that it is a reasonable approximation, in particular for processes that depend on electrostatic effects.[37–39]

The EVB/US Method

The above approaches for evaluation of the free-energy profile for ion transport may not be fully valid when applied to PTR processes, because in such cases the proton is likely to be transferred between water molecules rather than to be dragged along with water molecules. Thus, in studies of PTR, it is more effective to use the EVB method[40–43] that has been used very effectively in studies of PT in proteins in general and in the specific studies of the movement of protons across water chains in water and in proteins (e.g., Refs. 29, 44, 45; see also discussion and references in Ref. 23). The EVB has been described extensively elsewhere[41,42] and we will only describe it briefly for our specific case.

In studies of PTR between water sites, we describe the EVB quantum system in terms of diabatic states

$$\Psi_i = B_1B_2 \ldots B_iH^+ \ldots B_n \quad (5)$$

$$\Psi_j = B_1B_2 \ldots B_jH^+ \ldots B_n$$

where B_iH^+ is the protonated form of the B_i protonation site (e.g., a H_3O^+).

In the present case and in most other cases it has been found (e.g., Refs. 23, 46) that a two-state model allows one to capture the physics quite well when we have high activation barrier (which is of course the case in aquaporins). Thus, we use a two-state model with two water molecules and an excess proton as our EVB basis set:

$$\Psi_1 = H_3O^+_{(a)}H_2O_{(b)} \quad (6)$$

$$\Psi_2 = H_2O_{(a)}H_3O^+_{(b)}.$$

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 bond 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 these force-fields is given elsewhere[13]). The off diagonal elements of the Hamiltonian H_{12} is described by an empirical function that is fitted to experimental information and ab initio calculation, and the ground-state energy, E_g, is obtained by diagonalizing the EVB Hamiltonian, which gives for the two-state case:

$$E_g = \frac{1}{2}[(\varepsilon_1 + \varepsilon_2) - \sqrt{(\varepsilon_1 - \varepsilon_2)^2 + 4H_{12}^2}]. \quad (7)$$

The free-energy surface associated with the above E_g is obtained by the same FEP/US method mentioned above and changes the system adiabatically from state 1 to state 2 using

$$\varepsilon_m = (1 - \lambda_m)\varepsilon_1 + \lambda_m\varepsilon_2 \quad (8)$$

The free energy, ΔG_m, associated with changing λ_m from 0 to one is evaluated by the FEP procedure of Eq. 1. The free-energy functional that corresponds to the adiabatic ground state surface, E_g (Eq. 7), is then obtained by the FEP/US method as:

$$\Delta g(x') = \Delta G_m - \frac{1}{\beta}\ln\langle\delta(x - x')\exp[-\beta(E_g(x) - \varepsilon_m(x))]\rangle_m \quad (9)$$

where x' can be any type of coordinate that is used in mapping the free energy. In general, the reaction coordinate x' is defined by $\Delta\epsilon = \epsilon_1 - \epsilon_2$, 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 but projected in a subsequent step on the z-coordinate. λ_m is the λ that maximizes the chance to be in x.

More Standard PMF Methods

Another seemingly very effective way of evaluating the free energy of moving the ion from one position to another is the use of the US in the implementation introduced in our studies of electron transfer reactions[29,34] and in the evaluation of the free energies of the EVB approach (e.g., Ref. 32). In this approach, we use a mapping potential of the form:

$$\varepsilon_m = (1 - \lambda_m)\varepsilon_1 + \lambda_m\varepsilon_2 \quad (10)$$

$$\varepsilon_1(z) = E_g(z) + E_{1,cons}(z) = E_g(z) + K(z - z_0^{(1)})^2$$

$$\varepsilon_2(z) = E_g(z) + E_{2,cons}(z) = E_g(z) + K(z - z_0^{(2)})^2$$

where $z_0^{(1)}$ and $z_0^{(2)}$ are the two neighboring points on the z-axis and $E_g(z)$ is the total potential of the system (i.e., the solvated H_3O^+) without the constraints. We thus use our FEP/US formula[29,47]:

$$\Delta g(z) = -\frac{1}{\beta}\ln[\exp\{-\Delta G(\lambda_m)\beta\}\langle\delta(z' - z)\exp\{E_{m,\mathrm{cons}}(z)\beta\}\rangle_{\varepsilon_m}. \quad (11)$$

In using the PMF approach, one encounters frequently enormous hysteresis. This hysteresis is reduced here by evaluating the $\Delta G(\lambda_m)$ and the $\langle\delta(z' - z)\exp\{E_{m,cons}(z)\beta\}\rangle_{\epsilon_m}$ terms of Eq. 11 as the average of the forward and backward runs (see Ref. 47).

The PDLD/S-LRA Method

Despite the formal rigor of the FEP and LRA methods, it is found frequently that such methods are subjected to major convergence problems when one deals with electrostatic effects in protein interiors and that semimacroscopic models can give more reliable results. This is true in particular with regard to the PDLD/S-LRA method[31,48] 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:[48]

$$\Delta U_{\mathrm{sol},i}^{w\to p} = [-\Delta G_{\mathrm{sol},i}^{w} + \Delta G_{\mathrm{sol},p}^{w}(q = q_i) - \Delta G_{sol,p}^{w}(q = 0)] \times \left(\frac{1}{\varepsilon_p} - \frac{1}{\varepsilon_w}\right) + \Delta U_{q\mu}\frac{1}{\varepsilon_p} \quad (12)$$

where $\Delta G_{sol,i}^{w}$ is the free energy of solvation of the i-th ionizable group in water (the self-energy in water), $\Delta G_{sol,p}^{w}(q = q_i)$ and $\Delta G_{sol,p}^{w}(q = 0)$ are the free energies of solvation of the entire protein in water with atomic charges (q_i) present on the particular group ("charged state") and with atomic charges on the group set to zero ("uncharged state"), respectively. The $\Delta G_{sol,p}^{w}(q = 0)$ 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 than the actual protein dielectric constant (see Ref. 49).

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. Moreover, for accurate free-energy differences, several protein configurations should be averaged. The configurational space can be adequately sampled by using Monte Carlo or molecular dynamics (MD) techniques.[50] A closely related treatment has been developed by Alexov and Gunner.[51] In the present study, we used an MD approach in the LRA framework described above. This approach approximates the free energy associated with a transformation between two charged states by averaging the potential difference between the initial and final states over trajectories propagated on these two states, respectively. Using the PDLD/S free energy that corresponds to a single protein structure as an effective potential in the PDLD/S-LRA method, the free energy of solvation is given by:

$$\Delta G_{\mathrm{sol},i}^{w\to p} = \frac{1}{2}[\langle\Delta U_{\mathrm{sol},i}^{w\to p}\rangle_{q=q_1} + \langle\Delta U_{\mathrm{sol},i}^{w\to p}\rangle_{q=0}] \quad (13)$$

where the $\Delta U_{sol,i}^{w\to p}$ is the PDLD/S effective potential of Eq. 12, the $\langle\rangle_{q=q_1}$ 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 into account the reorganization of the environment explicitly, it may not fully account for some effects such as the complete water penetration and protein reorganization. These factors and the effect of induced dipoles are implicitly included in the model, which lead to the use of ϵ_p in the PDLD/S model.

The basic PDLD/S-LRA calculations are performed with all the protein groups (except some active site residues) in

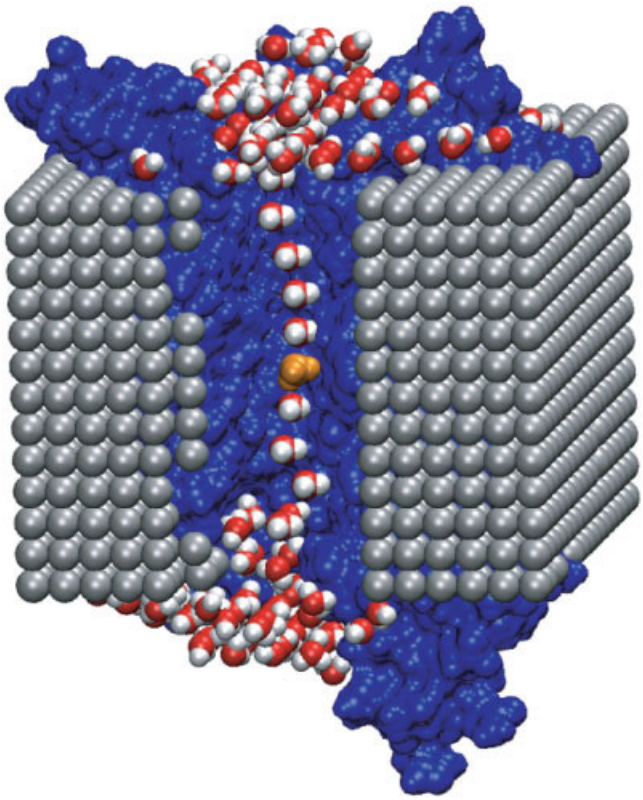

Fig. 3. The all-atom simulation system. The actual calculation involves the SCAAS spherical boundary conditions in which the surface constrained water sphere is shifted when the central H_3O^+ ion is shifted along the channel axis. The membrane region is modeled by a grid of 40 × 40 × 30 Å size and a 2.5 Å spacing of carbon-like atoms that represent the low dielectric membrane.

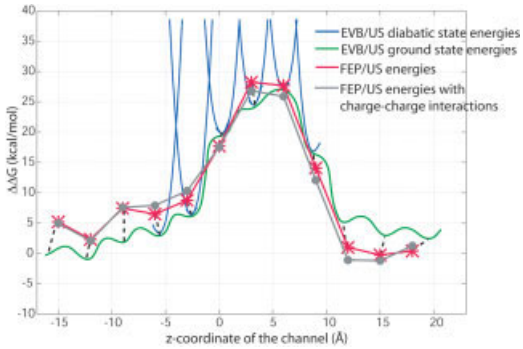

Fig. 4. Showing the EVB/US and FEP/US microscopic free energy profiles along the *z*-axis of the channel (the membrane starts and ends at $z = -14.0$ Å and $z = 8.5$ Å, respectively). FEP/US results were obtained with (grey) and without (red) the macroscopically estimated effect of the charge–charge interactions. EVB/US results are given both in terms of the diabatic free energies (blue) and the free energy of the adiabatic ground state (green) evaluated by Eq. 9.

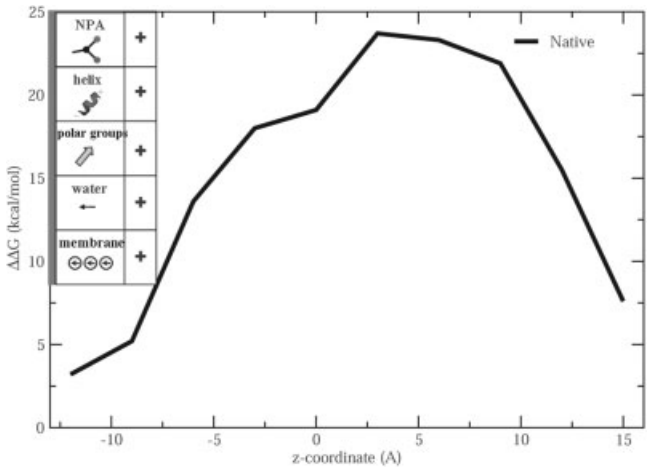

Fig. 5. The PDLD/S-LRA free-energy profile. The electrostatic elements included in the calculations are drawn schematically in the boxes on the left side of the figure.

their neutral form. The effect of ionizing these groups is evaluated macroscopically by finding their ionization state in a self-consistent way[49] and then evaluating the effect of these groups using $\Delta G_{ij} = 332 q_i q_j / (r_{ij} \epsilon_{ij})$, where r_{ij} is the distance between the interacting groups, and ϵ_{ij} is an effective dielectric constant whose value is determined by a distance-dependent function.[48,52] The justification of this approximation is discussed in detail elsewhere.[31,48,49] Basically, ϵ_{ij} for charge–charge interaction reflects the compensation of the gas phase Coulomb interaction between the charges by the solvation effect of the protein plus solvent system. This compensation has been found to be unexpectedly large even for charge–charge interaction in the protein interior, leading to a large effective ϵ_{ij} (between 20 to 40). This fact has been established repeatedly by both theoretical and experimental studies (e.g., Refs. 53, 54). It is also important to realize that ϵ_{ij} is not equal (and typically much larger) than the dielectric constant ϵ_p, that determines $\Delta\Delta G_{sol}^{w \to p}$ (see Ref. 49 and discussion below).

Simulation System

The present study was performed on *Escherichia coli* glycerol facilitator (GlpF) channel (Protein Data Bank entry: 1LDA) considering the simulation system shown in Figure 3. The construction of this system is described in Refs. 12 and 13, and here we only provide some key details.

The PDLD/S-LRA calculations involved as usual (e.g., Ref. 31) two steps, first running MD to generate protein configurations for the charged and uncharged states, and then averaging the PDLD/S results for the generated configurations. The MD runs were performed with the polarizable ENZYMIX force-field.[48] All the PDLD/S-LRA calculations were performed by the automated procedure of the MOLARIS program[48,55] where we generated typically 20 configurations for the charged and uncharged state, using MD simulations of 1 ps, with a 1-fs time step, for each configuration. The microscopic FEP/US and EVB/US calculations were performed using the ENZYMIX force-field,[48,55] with the solute parameters described in Ref. 13. The simulation included the use of 18 Å of the SCAAS spherical constraints and the local reaction field (LRF) long-range treatment.[56] The simulation system represented the membrane by a grid of induced dipoles (e.g., see Ref. 12) which were treated explicitly in our

polarizable model. The FEP/US simulation was done with 51 frames to transform the net charge of hydronium ion from 0 to +1, where each frame included 80 ps of simulation with 1-fs time steps. The EVB/US simulations of each proton transfer step were done with 34 frames, where each frame included 30 ps of simulation of the forward and backward runs for each. Convergence was improved by repeating the calculations from different initial conditions and averaging the resulting profiles. The runs from different initial conditions as well as longer runs help to establish our error range that we estimate to be approximately 2 kcal/mol.

RESULTS

As stated in the Introduction we will focus in this work on the effect of different factors on the electrostatic profile. This will be done by systematic "mutation" calculations. This step by step analysis will demonstrate that the barrier should not be attributed to specific residues but to the overall loss of solvation. We will also demonstrate that inconsistent calculations can lead to very different conclusions than those obtained by consistent calculations.

The Overall Profile

As a starting point, we evaluated the overall electrostatic barrier for the H_3O^+ ion through the GlpF channel. The microscopic calculations were performed by both the microscopic FEP/US and the EVB/US approaches, whereas the semimacroscopic study was performed by using the PDLD/S-LRA approach. The results of the microscopic calculations, which are depicted in Figure 4, indicated (as was found before[13] for AQP1) that there is a high barrier of about 27 kcal/mol in the center of the channel. The calculated barrier does not change significantly when the effect of the ionized groups is included (see also the section The Removal of All the Polar and Ionized Contributions Increases, Rather Than Decreases, the Barrier). As pointed out in our previous works,[12,13] and will also be demonstrated below, the high barrier is due to the fact that the "solvation" in the protein is significantly smaller than the corresponding solvation in water, because the sum of the different contributions does not provide as much stabilization as water does.

It is important to point out here that the results of the FEP/US involved the use of small constraints (5 kcal/mol/Å^2 on motions in the x, y directions). Calculations that did not involve any constraint produced a reduction of about 5 kcal/mol at 0 Å $> z >$ −7 Å because of the penetration of several water molecules that created a pocket around the H_3O^+ ion and solvated it from the x, y direction. Apparently, even in this case we obtained a high barrier at the center of the channel (about 25 kcal/mol). Fortunately, however, we are not dealing here with the PMF of moving the same H_3O^+ ion across the channel but with a PTR process where the proton moves from one water molecule to another. This means that we should consider the solvation of the H_3O^+ ion with constraint that retains an approximated chain of water molecules. This was done here (in the calculations summarized in Fig. 4) by introducing a 10 kacl/mol/Å^2 z-direction constraint on the H_3O^+ ion and the neighboring water molecules. In the EVB calculations, however, we did not need special distance constraint because the physics of the PTR process is inherent in the EVB treatment. We chose, however, to fix the oxygen of the donor and acceptor sites while calculating the free-energy profile of moving the proton between these sites. Obviously this approach is not completely rigorous because it introduces constraint on the reaction path. However, it is quite doubtful that one can obtain more consistent results by PMF approaches because of major hysteresis problems (see discussion in Ref. 47). Probably the profile at the top of the barrier can be slightly improved by releasing the constraint and running trajectories of the unrestraint EVB. The resulting effect can also be considered as a transmission factor, but it is nonetheless beyond the scope and aims of the present work.

We also attempted to evaluate the profile for H_3O^+ transfer by the regular PMF of Eq. 11. Unfortunately, this approach evolved enormous hysteresis perhaps reflecting the problems mentioned in our previous paper.[47] It is also possible that attempts to pull H_3O^+ through the channel are intrinsically problematic because of the very complex landscape of this system (where the solvation coordinate can change drastically upon penetration of more water molecules). Because we are interested in a PTR across a water chain, where the proton is the main transferred particle and the water chain just responds to the new proton positions, we did not find it useful to place more effort on the standard PMF calculations.

The PDLD/S-LRA electrostatic barrier is depicted in Figure 5, again giving a barrier of about 25 kcal/mol. It is significant to note, however, that the calculations presented here were obtained with $\epsilon_p = 2$ rather than $\epsilon_p = 4$ which has been used in most PDLD/S studies. The reasons for using this seemingly inconsistent treatment are in fact quite instructive. That is, as we note repeatedly (e.g., Ref. 49), ϵ_p represents all the effects that are not included explicitly in the given treatment. In principle, the PDLD/S includes all effects except the induced dipoles of the system and thus should involve the use of $\epsilon_p = 2$. However, in most cases, the LRA simulations do not reflect a fully converging treatment of the protein reorganization and in particular the effect of water permeation and thus require the use of $\epsilon_p = 4$. However, in the case of aquaporins we already have a chain of water molecule in the channel and thus the water permeation effect is taken into account in a more or less consistent way, so that $\epsilon_p = 2$ seems to be an optimal choice. In some respects, one can also argue that the PDLD/S-LRA dielectric has been calibrated by forcing it to reproduce the FEP/US barrier. It must be emphasized, however, that calculations that do not involve the LRA relaxation process would give very unrealistic results (see below).

At this point it might be useful to clarify that the EVB study of Ref. 18 gave also a high barrier (about 18 kcal/mol), whereas the PMF calculations of Ref. 11 gave (for the transfer from $z = -8$ Å to the NPA region) a barrier of only about 4 kcal/mol. This result seems to represent a

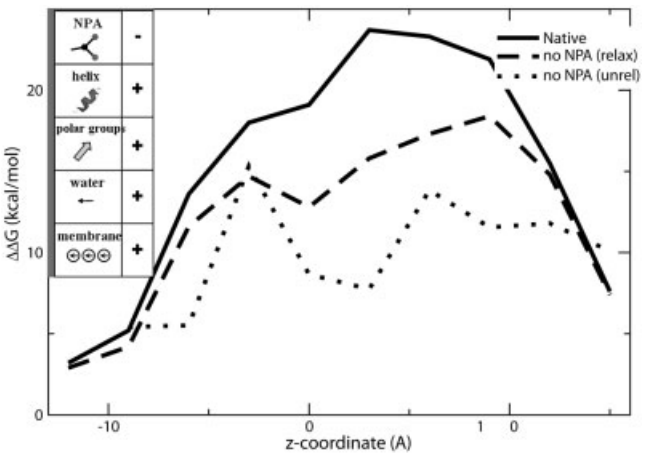

Fig. 6. The PDLD/S-LRA estimate of the effect of the NPA residues on the electrostatic barrier. The figure provides both the results for mutations of the NPA residues and the inconsistent estimates of the effect of the NPA motif obtained by the unrelaxed calculations using $\epsilon_p = 4$ (see text). The fact that the NPA region is mutated is designated by the minus sign in the corresponding box.

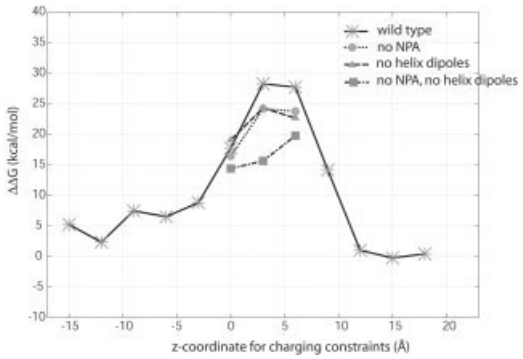

Fig. 7. The results of the FEP/US calculations of the effect of mutations of different electrostatic elements in aquaporin.

major underestimate that perhaps reflects the fact that the PMF was not evaluated relative to the bulk solvent. The continuum estimate of Ref. 11 is about 10 kcal/mol which also seems to represent a significant underestimate (it is difficult to see how such a barrier will lead to a blockage of the PTR process).

After obtaining the overall free-energy profile, we are ready to examine the different factors that contribute to the high overall barrier in a consistent way. We will also explore the dependence of the mutational effect on the method used and identify the origin of the different results by clear demonstration of the origin of various problems.

The Effect of Mutations of the NPA Motif

The examination of the contribution of the NPA motif to the barrier focused on Asn 68 and Asn 203 which constitute the main polar elements of this motif. This contribution was evaluated by both the PDLD/S-LRA and the microscopic FEP/US approach. The results of the PDLD/S-LRA calculations are reported in Figure 6. The FEP/US

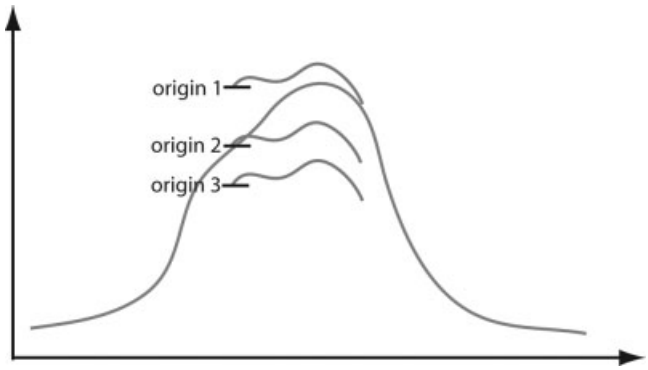

Fig. 8. Illustrating the problems with PMF calculations that do not use absolute reference in energy. The figure shows that shifting the origin of the PMF for a particular mutation up or down (from its height at the same point in the native enzyme) may lead to completely different effects of that mutation. Note that Ref. 11 used the same origin for all mutations instead of taking the energy on the bulk as the origin.

mutational results are summarized in Figure 7. As seen from the figure and in contrast to the perception of some workers (e.g., Refs. 4, 8), the NPA region does not lead to a large repulsion of the positive charge. This point can also be seen from the detailed analysis of Ref. 12 which reflects the fact that the $\Delta U_{q\mu}$ of Eq. 12 is reduced drastically upon relaxation.

In addition to the consistent PDLD/S-LRA calculations we tried to demonstrate the problem with the unrelaxed calculations. Our starting point is that treatments that do not allow the protein to relax would drastically overestimate the effect of the NPA motif (see Ref. 12). However, it is less obvious how to demonstrate the problems with the unrelaxed calculations in terms of mutational effects, since both the profiles of the native and the mutant are sometimes overestimated. Here we calculated first the native unrelaxed profile and found that using $\epsilon_p = 4$ instead of $\epsilon_p = 2$ makes the unrelaxed results similar to the relaxed results. Thus we considered the use of the unrelaxed mutant results with $\epsilon_p = 4$ as a way of presenting the inconsistent unrelaxed results and reported this profile in Figure 6. Other options for demonstrating the problems with the unrelaxed calculations include evaluation of the contribution from the mutant residue to the $\Delta U_{q\mu}$ term of Eq. 12, while using $\epsilon_p = 2$.

At any rate the unrelaxed mutational effect is probably related to the problematic conclusions obtained by other workers. For example, the PB results reported by Ref. 11 were obtained without relaxing the protein dipoles during the macroscopic charging process. Such a treatment may lead to the same type of overestimate reported in Ref. 12 and in Figure 6 for the unrelaxed model where the effect of the NPA residues is overestimated by up to 7 kcal/mol.

To verify the conclusions obtained from the PDLD/S-LRA calculations, it is important to examine the effect of mutating the NPA residues by microscopic calculations. This is done here by the FEP/US approach and the corresponding results are summarized in Figure 7. The calculations were done only at a few points because what counts is the change in the overall activation barrier. The

fact that the barrier is not reduced significantly is obviously in clear contrast to the proposal that the mutations would reduce the barrier in a major way. Indeed, as can be seen in Figure 7, the effect of the NPA mutation is quite small (only about −4 kcal/mol).

One may wonder at this point what is the difference between our microscopic calculations to the microscopic PMF calculations reported by Ref. 11. The answer seems to be associated with the fact that Ref. 11 has not provided a PMF for the actual transfer of the H_3O^+ from the bulk water to the channel center (see above). Thus, the relative height of the PMF profile was chosen arbitrarily (sometimes it was placed on the corresponding results of the continuum calculations without a justification). Now with such a treatment it is in fact impossible to determine the effect of the mutation in a quantitative way. This problem is illustrated schematically in Figure 8, where we show that different guesses of the initial height of the PMF will give different effects of mutations. Obviously, the above problem does not exist in our FEP/US approach because the height of each point of the PMF is determined in an absolute way relative to the corresponding solvation energy of H_3O^+ in water.

Interestingly, the effect of the NPA mutation was found to be rather small in the PB calculations of Ref. 11 (only about 2 kcal/mol). Similarly, a very small effect of mutating the NPA in AQP1 was found in Ref. 14

Mutating the Helix Macrodipole

An inspection of the structure of aquaporins[57] reveals two helices, M3 and M7, pointing their positive ends (which form the NPA motif) to the center of the channel (see Fig. 9). Thus, it is tempting to assume that the macrodipoles of these helices provide a major contribution to the electrostatic barrier. In fact, recent works[10,11] strongly support this view. At the outset, we point out that many proposals that considered the helix macrodipole as a major electrostatic factor (e.g., Refs. 58–64) were found to reflect an underestimate of the protein + solvent dielectric effect (see discussions in Ref. 65). Both experimental (e.g., Refs. 66, 67) and consistent theoretical studies (e.g., Ref. 65) have shown that the effect of the macrodipole is usually quiet small. Furthermore, it was also demonstrated[65] that the helix dipole effect manifests itself only by the microscopic dipoles of the few residues at the end of the helix (and thus it is simply a part of the general class of oriented microscopic dipoles).

To illustrate that the effect of the helices' macrodipole of aquaporin is not large, we performed both PDLD/S-LRA and FEP/US calculations which considered this effect in a consistent way. The results of the PDLD/S-LRA calculations are summarized in Figure 10, where we illustrate that the effect of the helix dipole is relatively small and that only calculations that keep the protein fixed during the charging process lead to the artifact of a large electrostatic contribution; the effect of the helix is only about 8 kcal/mol and it would have probably been smaller if we had allowed for more relaxation, whereas the artificial result of nonrelaxed calculations leads to a larger effect. To provide

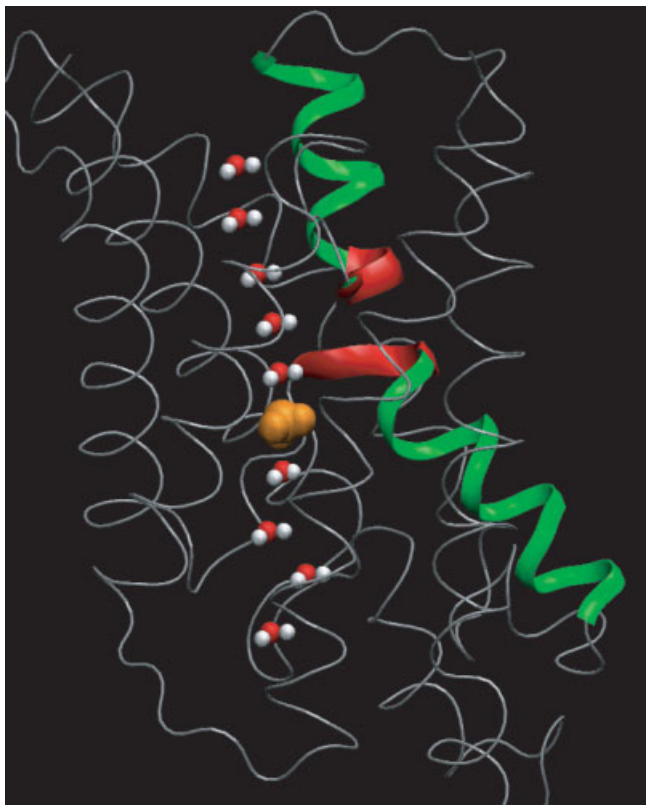

Fig. 9. Two helices that are assumed by some workers to be the origin of the barrier in aquaporin.

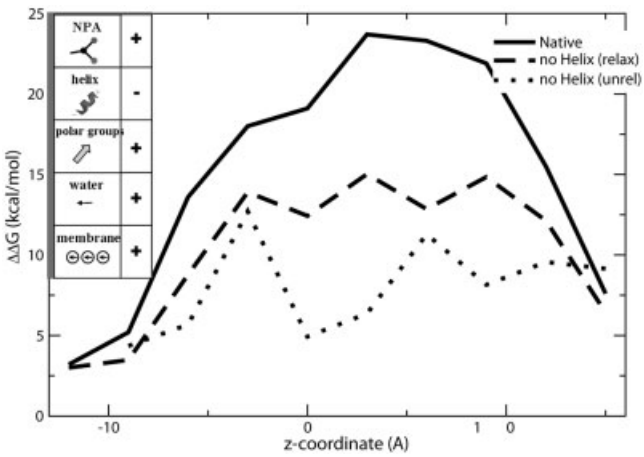

Fig. 10. The effect of the helix macrodipole on the electrostatic barrier. The figure depicts the consistent PDLD/S-LRA results as well as the inconsistent unrelaxed results.

a more complete picture, we also evaluate the combined effect of the NPA plus helix dipole (Fig. 11) and show that in contrast to recent suggestions[10] even this combined effect does not account for the barrier.

To further establish our findings, we also performed fully microscopic calculations of the effect of mutating the helix dipole. The results, which are summarized in Figure 7, illustrate that consistent calculations do not find a large effect of the helix macrodipole. We point out here again

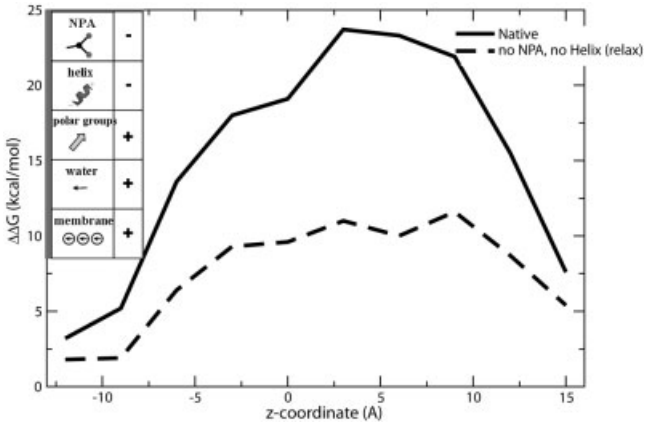

Fig. 11. The combined effect of the NPA and helix macrodipole.

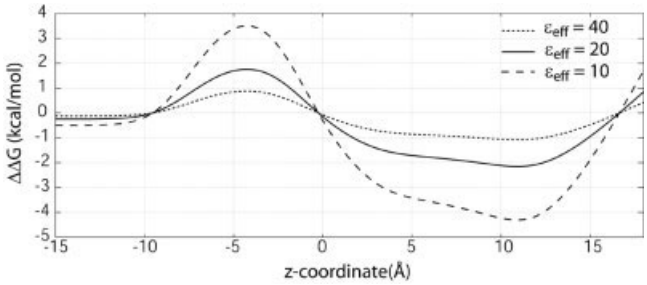

Fig. 12. The effect of the ionizable residues.

that the results of Ref. 11 that have been considered (e.g., Ref. 10) as evidence of a large effect of the helix dipole, probably reflect the above-mentioned problems. First, the PB calculations gave actually 3–4 kcal/mol effect of mutating the helix dipole. This is about 30% from the 10 kcal/mol total barrier estimated in Ref. 11 and also involves unrelaxed calculations. Second, the microscopic PMF calculations did not involve the essential evaluation of the absolute free energy at the initial point of the PMF profile (see Fig. 8). Furthermore, even the value obtained from the PMF calculations of Ref. 11 is about 4 kcal/mol, which does not seem to be a major contribution to the whole 25 kcal/mol barrier obtained in our studies that we consider to be a very realistic estimate (see section The Overall Profile). We note again that the PMF calculation did not provide an estimate of the overall barrier.

Mutating Ionized Residues

The effect of the protein ionizable group is clearly a factor that should be considered in examining the electrostatic barrier. However, as emphasized in many of our works [e.g., Refs. 49, 54 and in those of others (e.g., Ref. 53)], charge–charge interactions in proteins are smaller than what is usually assumed and correspond to having a large dielectric constant (e.g., Ref. 49). This point has already been emphasis in our early studies of aquaporins but we find it instructive to illustrate it again here. This is done in Figure 12, where we consider the effect of the protein ionized groups with dielectric constants of 40, 20, and 10 (where $\epsilon_{\text{eff}} = 10$ is more or less the lowest limit for the dielectric constant for charge–charge interaction). As seen from Figure 12, the overall effect of the ionizable groups is relatively small and cannot be considered as the origin of the electrostatic barrier. Note also (see section The Removal of All the Polar and Ionized Contributions Increases, Rather Than Decreases, the Barrier) that if all the ionizable groups were removed in a hypothetical hydrophobic channel, we would end up with a very high desolvation barrier.

To further clarify our point, we also calculated microscopically the effect of mutating Arg 206 and found it to shift the profile by 11, 15, and 12 kcal/mol for z values of -5, -3, and 0 Å. This large effect, which might also be easily reproduced by others with microscopic models, is clearly attributable to insufficient relaxation of the model and to the relatively short simulation time. This type of problem (which might look surprising to those who are not familiar with validation studies of charge–charge interactions) is often observed in clear benchmarks, where the actual charge–charge interaction is known (e.g., Ref. 53). Of course, removing selectively several positively charged groups might reduce the barrier significantly if the channel would not collapse or unfold. However, we are talking here on the overall effect of the positive and negative ionized groups on the barrier in the native enzyme (this effect can be assessed experimentally only by removing all positively and negatively charged ionized groups).

The Removal of All the Polar and Ionized Contributions Increases, Rather Than Decreases, the Barrier

Finally, we can move to an important point that has been perhaps misunderstood by several workers. That is, as pointed out by BW,[13] once we mutate all the protein atoms to nonpolar atoms we will end up with a nonpolar pore, which will in fact lead to a very high barrier. Pointing out this fact, which is illustrated in Figure 13, does not constitute a proposal that aquaporin is a low dielectric system (as was presumed by some, e.g., Ref. 10). It only establishes the point that the combined mutations of all the residues to their nonpolar form leads to a reduction of the barrier (relative to the fully nonpolar protein) rather than to the formation of a barrier for PTR (as is assumed by those who attribute the barrier to the polar groups). This point is not a semantic point but a fundamental issue in the description of the physics of the proton blockage by aquaporin. That is, the loss of solvation of H_3O^+ upon moving from water to protein is the general case rather than a specific case and the problem for evolution is not to create a barrier but to reduce the barrier when this is needed for a specific function.

It is also useful to point out that a fully nonpolar aquaporin will resemble a nanotube, where again, in contrast to the assumption in several recent works, the nonpolar pore will create a very large barrier for PTR (see discussion in Ref. 68).

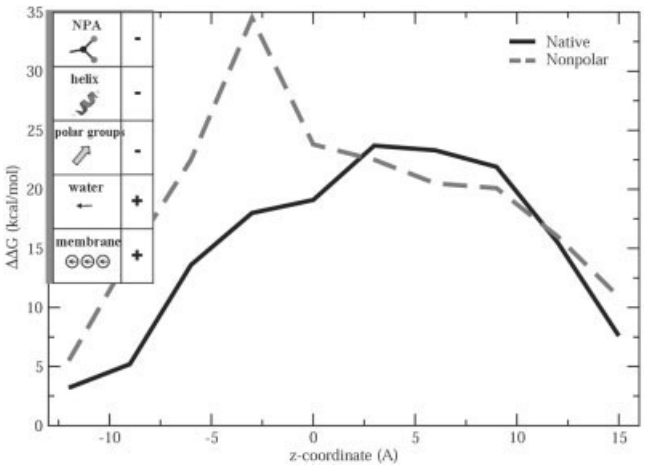

Fig. 13. The effect of mutating all the protein residual charges to zero.

TABLE I. Calculated and Observed Solvation Free-Energies and pK_as in Some Test Cases†

	Lysozyme		Staphylococcal nuclease	
System residue	Asp 52	Glu 35	Glu 66 (original)	Glu 66 ($Q = -1.68$)
FEP				
$\Delta G^{p}_{sol}(A^-)$	−80.5	−77.3	−47	−58
$\Delta G^{w}_{sol}(A^-)$	−79.5	−78.7	−78.7	−78.7
$\Delta\Delta G^{w\rightarrow p}_{sol}(A^-)$	−0.9	1.3	32	21
$\Delta\Delta G^{w\rightarrow p}_{sol}(AH)$	0.3	0.2	6	5
$\Delta\Delta G_{charges}$[a]	0.2	1.1	0.1	0.1
pK_{app}	3.4	5.7	19	11.4
PDLD/S-LRA				
pK_{app}	4.0	6.4	9.4	
Obs				
pK_{app}	3.5–3.7	6.1	8.8	8.8

† Free energy in kcal/mol and pK_as in pK_a units.
[a] $\Delta\Delta G_{charges}$ reflects the effect of other ionizable residues.

VALIDATING THE SIMULATION APPROACHES IN CASES IN WHICH THE EFFECTS OF MUTATIONS ON pK_as ARE KNOWN

As indicated above, the key requirement for a reliable evaluation of the free-energy profile for PTR through aquaporin is the ability to evaluate the solvation energy of H_3O^+ in different sites of the protein and in water [The $\Delta G(V_a \rightarrow V_b)$ of Eq. 2]. This solvation free energy is directly related to the pK_a of the H_3O^+ ion because[69]

$$pK^{p}_{a} = pK^{w}_{a} - \delta_0 \Delta\Delta G^{w\rightarrow p}_{sol}/(2.3RT) \qquad (14)$$

where ΔG_{sol} is the change in solvation energy upon moving from the un-ionized to the ionized form of the given group and $\Delta\Delta G^{w\rightarrow p}_{sol}$ is the change in ΔG_{sol} upon moving the group from solution to its site in the protein. The factor δ_0 is plus or minus for basic and acidic groups, respectively. Thus, the reliability of the calculated profile is directly related to the reliability of the corresponding pK_a calculations. Because we established above that some approaches do not provide a reliable tool for studies of the electrostatic barrier in aquaporin, we have to illustrate the reliability of our approach in benchmarks in which the experimental pK_as are known. The validity of the PDLD/S-LRA model has already been examined with the help of a discriminative benchmark in Ref. 49. This work also demonstrated the enormous problem with macroscopic models that use a low dielectric and do not allow the protein to relax. Thus we will focus here on the FEP/US calculation. In doing so, we chose the pK_a of Asp 52 and Glu 35 in lysozyme (e.g., see Ref. 31). We emphasize that, in contrast to the case with semimacroscopic studies, microscopic FEP calculations can be validated even by calculations of the pK_as of surface groups. That is, in case of semimacroscopic models it is essential to consider a benchmark of groups with a large pK_a shift, because studies of surface groups always give a small shift relative to water which is the observed results, whereas FEP calculations involve evaluation of absolute solvation free energy and there is no guarantee that the result will be correct at any environment (including on the surface of the protein). At any rate, the results for lysozyme are summarized in Table I, where we also provide the PDLD/S-LRA results. Both results are in a very good agreement with the corresponding experimental results.

To gain further insight, we also tried to examine the very challenging problem of the pK_a of Glu 66 in the hydrophobic site of staphylococcal nuclease (SNase). Reproducing the observed effect has been found to present a major challenge to macroscopic models.[70] Interestingly, our standard attempt to obtain the pK_a of Glu 66 by the FEP approach gave, as expected, a major deviation from the observed pK_a (predicted shift of 23 units instead of 5) despite running long simulations of 5 ns. This indicates that the protein undergoes a local unfolding and/or water penetration process upon ionization of Glu 66. The problem is, of course, to capture this configurational change by the available simulation times (e.g., see Ref. 71). To address this challenge, we introduced here a new approach in which the conformational change is induced by the charging coordinate, decreasing the charge from $Q = -1.0$ to $Q = Q_i$ where $|Q_i| > 1.0$. Performing such a charging and uncharging cycle with $Q_i = -2.0$ led to the remarkable hysteresis that can be seen by inspection of Figure 14. Thus, we do not present in Figure 14 the complete cycle but rather uncharged the group from −1.0 to 0.0 starting with the structure obtained at the final step of the full charging process (with the specific Q_i). Examining the structures from different charging cycles showed that charging up to $Q = Q_i = -2.0$ leads to a drastic unfolding and a complete exposure of the ionized Glu residue to the solvent. However, a gradual cycle that reached $Q = -1.7$ led to a small local change and water penetration. Because our overcharging procedure can generate different sets of protein structures (that correspond to different degrees of charged induced reorganization), we can use these structures to explore the relationship between the protein landscape and the energetic of the ionized acids. Our actual procedure is described elsewhere[72] and we only note here that we consider the cycle:

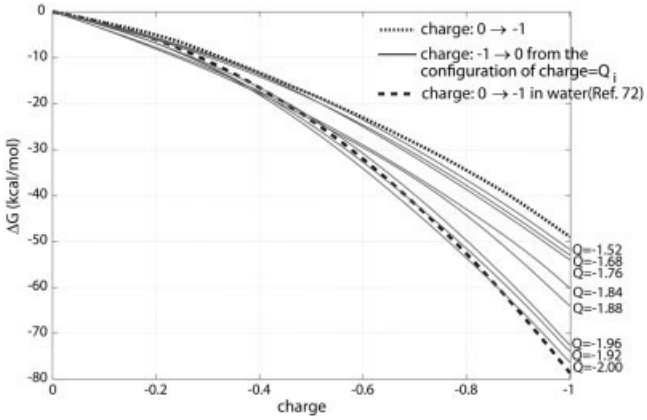

Fig. 14. The hysteresis for uncharging Glu 66 in the hydrophobic site of SNase from $Q = -1$ to $Q = 0$, starting from protein structures ($\mathbf{r}_i$) generated by overcharging from $Q = 0$ to $Q = -1$. Our actual procedure for evaluating the pK_a of Glu 66 involves the use of the cycle of Eq. 15 (see description in Ref. 72), but the present figure should help in clarifying the fact that different protein structures lead to different stabilization of the ionized form of Glu 66.

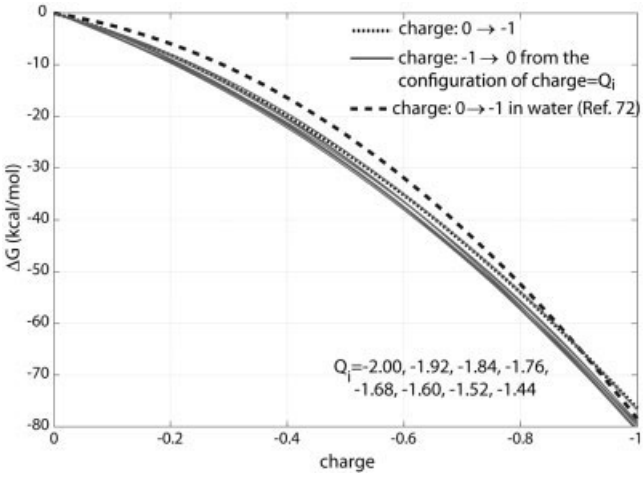

Fig. 15. The effect of using configurations generated by the overcharging process, in the case of Glu 35 in lysozyme. In contrast to the case of Glu 66 in SNase, we do not have here a significant hysteresis.

$$\Delta G(Q = 0.0 \rightarrow Q = -1.0) = \min[\Delta G(\mathbf{r}_0(Q = 0) \rightarrow \mathbf{r}_i(Q = Q_i)) + \Delta G(\mathbf{r}_i(Q = Q_i) \rightarrow \mathbf{r}_i(Q = -1.0))] \quad (15)$$

where $\mathbf{r}$ is the protein structure generated by charging up to Q and the first term in the right-hand side of Eq. 15 is evaluated by calculating the PMF for moving from $\mathbf{r} = \mathbf{r}_0$ and $Q = 0.0$ to $\mathbf{r} = \mathbf{r}_i$ and $Q = Q_i$. The second term in the right-hand side of Eq. 15 is obtained by calculating the PMF for moving with $\mathbf{r}$ fixed as $\mathbf{r} = \mathbf{r}_i$ from $Q = Q_i$ to $Q = -1.0$. The results obtained by our overcharging procedure are summarized in Table I. Obviously, the present findings require further validation but they clearly indicate that water penetration may have a major role in leading to the observed pK_a of Glu 66 (see Ref. 73 for such proposal). We also point out that the use of the artificial charging process as a way of simulating a physically meaningful water penetration may present an exciting new strategy in electrostatic calculations.

Because the overcharging approach gave a large hysteresis for Glu 66 in SNase, it is justifiable to ask whether this approach will not change the results reported above for lysozyme. To explore this issue, we repeated the procedure of Figure 14 for Glu 35 in lysozyme. As seen from Figure 15, we did not obtain significant hysteresis and we basically retaining the results of Table I, even after going to $Q = -2.0$ and using the resulting structure in the uncharging from $Q = -1.0$. Although many more studies are needed for exploring our approach, it seems that the overcharging approach allows us to sample the water penetration process and to see if it leads to a major change in the ionization process.

Despite the interesting insight obtained from the overcharge approach, we did not try to apply it to aquaporin, because of what we see as a time scale issue. That is, it is possible that there are very slow fluctuations that involve major water penetration and perhaps partial unfolding in aquaporin, but these slow effects are not the subject of the present article. More specifically, because the present article focuses on the effect of mutations in folded aquaporins and on the misunderstandings of previous electrostatic considerations, we do not find it useful to complicate the discussion of the electrostatic barrier in aquaporin by using the overcharging analysis. Thus, we can only say that our state of the art approaches were validated by pK_a calculations in lysozyme and that the same approaches with the overcharging strategy can be used to reproduce the pK_a of Glu 66 in SNase and other biological charge transfer processes that involve a full equilibration. However, the application of the overcharging approach to the possible nonequilibrium problem in aquaporin is left to subsequent studies.

CONCLUSIONS

The central point of this article has been the examination of the origin of the barrier for the transport of protons through the aquaporin channel. Accepting that the barrier is controlled by the corresponding electrostatic energy (see below), we quantified the origin of this barrier by careful computational mutations and examined the reason for the diverse results obtained by different studies. We also used this system as a powerful example of the problems in different treatments of electrostatic energies in proteins. It was found that the barrier is due to the fact that the "solvation" of the proton charge is smaller in the channel than in water. It should be clarified in this respect that the solvation of the charge reflects all the electrostatic components including the effect of the protein polar groups and ionized residues.[36]

The first step of our analysis focused on the overall height of the barrier for the PTR process. Our FEP/US and EVB/US approaches provide a barrier of about 25 kcal/mol that seems to reflect consistent and converging microscopic analysis. It is important to mention in this respect that other workers also obtained a relatively high barrier (e.g., 18 kcal/mol in Ref. 18).

Despite the uncertainties in the calculated PTR profile, it is quite clear that the barrier estimate is reasonable and that the very small barrier obtained by the partial PMF of Ref. 11 or even the PB barrier of that work is probably an underestimate. It is also important to mention that our treatment involved the explicit effect of the membrane-induced dipoles and the exclusion of water by the membrane region. The resulting membrane effect leads to significant increase in the barrier. However, our main point of this work is not the exact height and/or shape of the barrier but the nature of the contributions to this barrier.

Unfortunately, the calculations of the height of the barrier and even the evaluation of the electrostatic field from the protein do not provide any solid clue about the residues that contribute this barrier. In fact, as we emphasized before (e.g., Ref. 13), the specific contributions cannot be determined computationally by just evaluating the electrostatic interaction at a fixed protein structure. Such effects can be evaluated qualitatively by "mutating" the given residue with either the microscopic FEP/US or the PDLD/S-LRA approaches. Both approaches indicated that the polar groups of the NPA region do not lead to a large contribution to the electrostatic barrier. It was also demonstrated that the neglect of the relaxation of the system upon mutation of the NPA residues can lead to a major overestimate of the effect of these residues.

Careful and consistent examinations of the effect of the helix macrodipole indicated that this electrostatic element does not contribute to the penetration barrier in a major way. Because this finding seems to be in clear contrast to the suggestions of several workers (e.g., Refs. 8, 10, 11), we analyzed the origin of these discrepancies. This was done by demonstrating that the inconsistent use of macroscopic models with a low dielectric constant and without allowing the protein to relax would result in major overestimates of the effects of the helix macrodipole. The same problem seems to exist in PMF calculations that do not asses the free energy of moving the proton from the bulk solvent to the given protein site (thus producing a small barrier and relatively large effects of the helix dipoles). It may be useful to point out here that our prediction of the small effect of the helix dipole can be easily verified by placing a negative residue at the end of the helix and "neutralizing" the helix dipole. We predict that there will not be a significant change in the proton blockage upon such mutations. In fact, our early related prediction[65] has been confirmed experimentally in the case of triosephosphate isomerse.[74] Furthermore, our prediction[75] of the small effect of the Helix dipole in the KscA channel has been just confirmed experimentally.[76]

Our studies also pointed out that the combined effect of the protein ionized groups is relatively small. We also established that a fully nonpolar protein gives an even higher barrier than the barrier obtained while including all the polar contribution, that supposedly lead to the barrier.

The findings of the present work reproduced and quantified our early conclusions[12,13] and establish that the electrostatic barrier is attributable to the loss of the generalized solvation upon moving H_3O^+ ion from the bulk solvent to the center of the protein. It is crucial to reclarify in this respect that (in contrast to the possible impression from Ref. 10) it has not been suggested in Ref. 13 that the penetration barrier is due to the "fact" that aquaporin is a low dielectric system [Fig. 2(b)]. In fact, the studies of Refs. 12 and 13 considered all the electrostatic contributions in aquaporin [Fig. 2(d)] and found that the overall generalized solvation provided by these elements contributes much less generalized solvation than that of the corresponding stabilization of the H_3O^+ ion in bulk water.

The discussion of the biological significance of the NPA motif and/or the helix dipole (e.g., Ref. 10) seems to overlook several important points about what is required from an ion (or proton channel). In general it is not hard to create a channel that will not conduct protons. All that is needed is a relatively hydrophobic interior. Even a mildly polar interior is not sufficient for allowing proton transfer because the protons will be less stable than in water. Similarly it is not difficult to design a channel that will allow a moderate transport of water because the solvation energy of water is not large. However, the construction of a narrow channel that has the correct polarity to allow PTR may be a significant challenge for protein design.

Although the present article did not focus on the early ideas (e.g., Refs. 5, 9) that the barrier in aquaporin is due to water orientational effects that disrupt the Grotthuss mechanism, it is useful to make here some related comments: Careful considerations and actual simulations have demonstrated (e.g., Ref. 46) that the results obtained with a complete EVB treatment (that includes the proton and a significant number of water molecules in the EVB Hamiltonian) give penetration barriers that are similar to those obtained with the simplified Marcus treatment or with the two-state EVB/US approach used here. In other words, as long as we have a relatively high solvation barrier, which is extended on more than two water molecules (so a concerted motion cannot reduce the barrier), we will obtain similar barriers either by taking into account all the quantum mechanical delocalization effects or by just evaluating the solvation free energy of the localized H_3O^+ site.

Apparently the nature of the modified Marcus model or the simplified EVB model has not been fully appreciated by some workers. For example, Ref. 14 referred to the simplified EVB as an empirical model, which is parameterized on bulk water and thus presumably cannot be used in the aquaporin environment. This work also considered the two-state EVB model as a problematic model because it is presumably based on the solvation of H_3O^+ rather than the solvation of $H_5O_2^+$. A related misunderstanding was expressed by Ref. 11 where it was assumed that Ref. 13 did not consider delocalization effects. This assumption seems to reflect unfamiliarity with the ability of the EVB model to describe PTR processes. That is, the full EVB and the simplified EVB use as a basis set the energetics of the H_3O^+ states (Eq. 5) but then mix these states in a consistent way and obtain the delocalization effects, reproducing $H_5O_2^+$ and other delocalized configurations (this is

done, of course, while incorporating the effects of different environments). However, starting with gas phase calculations of the $H_5O_2^+$ system and then solvating this system is a fundamentally inconsistent approach (with incorrect solute–solvent coupling), which is similar in many respects to the QM-FEP approach,[77] whose problems (relative to consistent QM/MM treatments) have been discussed elsewhere (e.g., Ref. 78). At any rate, we believe that the simplified EVB provides an excellent tool for capturing the origin of the penetration barrier and that an EVB treatment with more extended delocalization will not change our conclusion significantly. This point has been established quantitatively in the test case of gramicidin.[46] Finally, we point out that even if the proton charge were delocalized on three water molecules at the top of the barrier, we would get basically the same barrier as that obtained by the localized model.

At this stage, it may be instructive to summarize and reclarify the key points of the present article: i) The overall barrier for PTR in proteins in general and in aquaporin in particular is determined by the overwhelming desolvation barrier and this can be modulated by specific electrostatic interactions. In fact, there is no need in any specific electrostatic interaction to create a barrier for PTR. ii) The only way to examine whether specific interactions contribute to the barrier is to mutate (either experimentally or theoretically) the corresponding residues and to look at the location of the highest point on the electrostatic profile. iii) To the best of our knowledge, the earliest attempt to actually determine the importance of specific electrostatic interactions was reported by BW,[12] who included in their study all the relevant electrostatic interactions and treated them in a consistent way. These workers found that the barrier is primarily attributable to the inability of the protein to provide sufficient solvation to the H_3O^+ ion. iv) The finding of BW has been confirmed and quantified by the studies of Ref. 12 and by the present work. v) BW have never suggested that aquaporin can be viewed as a nonpolar pore and in fact these workers estimated an effective dielectric of about 7 for the H_3O^+ penetration process. vi) Subsequent studies that attempted to estimate the effect of different electrostatic elements have seemed to be problematic, reflecting either the neglect of the protein reorganization or a PMF with an arbitrary reference point.

In view of the above considerations, we predict that mutations of the NPA region will lead to a relatively small increase in proton conductance through the aquaporin barrier. We also clarify once more what must be meant by the argument that the barrier is "due" to the NPA motif. Such a proposal means that a removal of the NPA will reduce the barrier in a drastic way. Because it was found here that the removal of the NPA does not remove the barrier; the barrier is not due to this motif. The same statement is also applied to the helix dipole. The barrier actually reflects the fact that the overall generalized solvation energy provided to the H_3O^+ ion by the channel is less than that provided by the environment of an H_3O^+ in the bulk water.

It might be useful to reiterate here our point about the contributions that lead to the barrier. That is, if we define contributions by mutating the charges or dipoles of the corresponding elements, we find that some mutations lead to a small but not negligible effect. Still, all of these effects are significantly smaller than the overall barrier. Thus, these contributions cannot be the origin of the barrier. Furthermore, mutating all the residual charges still leaves us with a very high barrier so that the sum of all the effects of the protein charges cannot be the origin for the barrier. The point that may have been overlooked by some workers is the fact that the barrier of moving the H_3O^+ to a nonpolar protein (*which is not our model*) is so high that adding polar elements usually pushes the barrier down even if the unrelaxed configuration of the polar residues destabilizes the ion. Thus, the barrier is due to the *change* in the solvation energy of the ion, which in its general form[79] includes the interaction with the protein polar groups. It is also important to emphasize that this was already the conclusion of our first article.[13]

Although this article seems to send a "negative" message with regard to the predicted success of mutation experiments (see related prediction in Ref. 80), it does not mean that the situation is hopeless. In principle, if it would be possible to keep the channel folded while performing arbitrary mutations, it should be possible to increase the proton conductance. For example, lining the NPA region with serines would have drastically reduced the barrier at this point. Of course, the challenge is to reduce the barrier at all the high-energy regions.

ACKNOWLEDGMENTS

The authors thank the University of Southern California's High-Performance Computing and Communications (HPCC) Center for computer time.

REFERENCES

1. Agre P, Kozono D. Aquaporin water channels: molecular mechanisms for human diseases. FEBS Lett 2003;555(1):72–78.
2. Murata K, Mitsuoka K, Hiral T, et al. Structural determinants of water permeation through aquaporin-1. Nature 2000;407:599–605.
3. Sui H, Han B-G, Lee JK, Walian P, Jap BK. Structural basis of water-specific transport through the AQP1 water channel. Nature 2001;414:872–878.
4. Yarnell A. Blockade in the cell's waterway. Chem Eng News 2004;82(4):42–44.
5. de Groot B, Grubmuller H. Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. Science 2001;294:2353–2357.
6. de Groot BL, Frigato T, Helms V, Grubmuller H. The mechanism of proton exclusion in the aquaporin-1 water channel. J Mol Biol 2003;333(2):279–293.
7. Decoursey TE. Voltage-gated proton channels and other proton transfer pathways. Physiol Rev 2003;83:475–579.
8. Jensen MO, Tajkhorshid E, Schulten K. Electrostatic tuning of permeation and selectivity in aquaporin water channels. Biophys J 2003;85(5):2884–2899.
9. Tajkhorshid E, Nollert P, Jensen M, Miercke L, Stroud RM, Schulten K. Control of the selectivity of the aquaporin water channel by global orientational tuning. Science 2002;296:525–530.
10. de Groot BL, Grubmuller H. The dynamics and energetics of water permeation and proton exclusion in aquaporins. Curr Opin Struct Biol 2005;15(2):176–183.
11. Chakrabarti N, Roux B, Pomes R. Structural determinants of proton blockage in aquaporins. J Mol Biol 2004;343(2):493–510.

12. Burykin A, Warshel A. On the origin of the electrostatic barrier for proton transport in aquaporin. FEBS Lett 2004;570(1–3):41–46.
13. Burykin A, Warshel A. What really prevents proton transport through aquaporin? Charge self-energy versus proton wire proposals. Biophys J 2003;85:3696–3706.
14. Miloshevsky GV, Jordan PC. Water and ion permeation in bAQP1 and GlpF channels: a kinetic Monte Carlo study. Biophys J 2004;87(6):3690–3702.
15. Agmon N. The Grotthuss mechanism. Chem Phys Lett 1995;244: 456–462.
16. Eigen M. Proton transfer, acid-base catalysis, and enzymatic hydrolysis. Part I: Elementary processes. Angew Chem Int Ed Engl 1964;3(1).
17. Zundel G, Fritcsh J. The chemical physics of solvation. Vol. 2. Amsterdam: Elsevier; 1986. Chapter 2.
18. Ilan B, Tajkhorshid E, Schulten K, Voth GA. The mechanism of proton exclusion in aquaporin channels. Proteins 2004;55(2):223–228.
19. Sham Y, Muegge I, Warshel A. Simulating proton translocations in proteins: probing proton transfer pathways in the *Rhodobacter sphaeroides* reaction center. Proteins 1999;36:484–500.
20. Warshel A. Conversion of light energy to electrostatic energy in the proton pump of *Halobacterium halobium*. Photochem Photobiol 1979;30:285–290.
21. Warshel A, Åqvist J, Creighton S. Enzymes work by solvation substitution rather than by desolvation. Proc Natl Acad Sci USA 1989;86:5820–5824.
22. Warshel A, Ottolenghi M. Kinetic and spectroscopic effects of protein-chromophore electrostatic interactions in bacteriorhodopsin. Photochem Photobiol 1979;30:291.
23. Braun-Sand S, Strajbl M, Warshel A. Studies of proton translocations in biological systems: simulating proton transport in carbonic anhydrase by EVB-based models. Biophys J 2004;87(4):2221–2239.
24. Warshel A, Sussman F, King G. Free energy of charges in solvated proteins: microscopic calculations using a reversible charging process. Biochemistry 1986;25:8368–8372.
25. Lee FS, Chu ZT, Bolger MB, Warshel A. Calculations of antibody-antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603. Protein Eng 1992;5:215–228.
26. Sham YY, Chu ZT, Tao H, Warshel A. Examining methods for calculations of binding free energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA calculations of ligands binding to an HIV protease. Proteins 2000;39:393–407.
27. Zwanzig RW. High-temperature equation of state by a perturbation method. I. Nonpolar gases. J Chem Phys 1954;22:1420.
28. Valleau JP, Torrie GM. A guide to Monte Carlo for statistical mechanics: 2 Byways. In: Berne BJ, editor. Modern theoretical chemistry. Vol. 5. New York: Plenum Press; 1977. p 169–194.
29. Warshel A. Dynamics of reactions in polar solvents. Semiclassical trajectory studies of electron-transfer and proton-transfer reactions. J Phys Chem 1982;86:2218–2224.
30. Warshel A. Calculations of chemical processes in solutions. J Phys Chem 1979;83:1640–1650.
31. Sham YY, Chu ZT, Warshel A. Consistent calculations of pK_a's of ionizable residues in proteins: semi-microscopic and macroscopic approaches. J Phys Chem B 1997;101:4458–4472.
32. Hwang J-K, King G, Creighton S, Warshel A. Simulation of free energy relationships and dynamics of S_N2 reactions in aqueous solution. J Am Chem Soc 1988;110(16):5297–5311.
33. Åqvist J, Warshel A. Energetics of ion permeation through membrane channels. Solvation of Na^+ by gramicidin A. Biophys J 1989;56:171–182.
34. King G, Warshel A. Investigation of the free energy functions for electron transfer reactions. J Chem Phys 1990;93:8682–8692.
35. Kubo R, Toda M, Hashitsume N. Statistical physics II: nonequilibrium statistical mechanics. Berlin: Springer-Verlag; 1985.
36. Warshel A, Russell ST. Calculations of electrostatic interactions in biological systems and in solutions. Q Rev Biophys 1984;17:283–421.
37. Hwang J-K, Warshel A. Microscopic examination of free energy relationships for electron transfer in polar solvents. J Am Chem Soc 1987;109:715–720.
38. Kuharski RA, Bader JS, Chandler D, Sprik M, Klein ML, Impey RW. Molecular model for aqueous ferrous ferric electron transfer. J Chem Phys 1988;89:3248–3257.
39. Åqvist J, Hansson T. On the validity of electrostatic linear response in polar solvents. J Phys Chem 1996;100:9512–9521.
40. Warshel A, Weiss RM. Energetics of heme-protein interactions in hemoglobin. J Am Chem Soc 1981;103:446.
41. Warshel A. Computer modeling of chemical reactions in enzymes and solutions. New York: John Wiley & Sons; 1991.
42. Aqvist J, Warshel A. Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. Chem Rev 1993;93:2523–2544.
43. Warshel A. Computer simulations of enzyme catalysis: methods, progress, and insights. Annu Rev Biophys Biomol Struct 2003;32: 425–443.
44. Schmitt UW, Voth GA. Multistate empirical valence bond model for proton transport in water. J Phys Chem B 1998;102:5547–5551.
45. Vuilleumier R, Borgis D. An extended empirical valence bond model for describing proton transfer in $H^+(H_2O)_n$ clusters and liquid water. Chem Phys Lett 1998;284:71–77.
46. Braun-Sand S, Burykin A, Chu ZT, Warshel A. Realistic simulations of proton transport along the gramicidin channel: demonstrating the importance of solvation effects. J Phys Chem B 2005;109: 583–592.
47. Kato M, Warshel A. Through the channel and around the channel: validating and comparing microscopic approaches for evaluation of free energy profiles for ion penetration through ion channels. J Phys Chem B 2005;109:19516–19522.
48. Lee FS, Chu ZT, Warshel A. Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. J Comp Chem 1993;14:161–185.
49. Schutz CN, Warshel A. What are the dielectric "constants" of proteins and how to validate electrostatic models. Proteins 2001; 44:400–417.
50. Allen MP, Tildesley DJ. Computer simulation of liquids. Oxford: Oxford University Press; 1987.
51. Alexov E, Gunner MR. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. Biophys J 1997;72:2075.
52. Warshel A, Russell ST, Churg AK. Macroscopic models for studies of electrostatic interactions in proteins: limitations and applicability. Proc Natl Acad Sci USA 1984;81:4785–4789.
53. Johnson ET, Parson WW. Electrostatic interactions in an integral membrane protein. Biochemistry 2002;41:6483–6494.
54. Sham YY, Muegge I, Warshel A. The effect of protein relaxation on charge-charge interactions and dielectric constants of proteins. Biophys J 1998;74:1744–1753.
55. Chu ZT, Villa J, Strajbl M, Schutz CN, Shurki A, Warshel A. MOLARIS version beta9.05. 2004.
56. Lee FS, Warshel A. A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. J Chem Phys 1992;97:3100–3107.
57. Fu D, Libson A, Miercke L, Weitzmann C, Nollert P, Krucinski J, Stroud RM. Structure of a glycerol conducting channel and the basis for its selectivity. Science 2000;290:481–486.
58. Hol WGJ, Duijnen PTV, Berendson HJC. Alpha helix dipole and the properties of proteins. Nature 1978;273:443–446.
59. Wada A. The α-helix as an electric macro-dipole. Adv Biophys 1976;9:1–63.
60. Daggett VD, Kollman PA, Kuntz ID. Free-energy perturbation calculations of charge interactions with the helix dipole. Chem Scripta 1989;29A:205.
61. Gilson M, Honig B. The energetics of charge-charge interactions in proteins. Proteins 1988;3:32–52.
62. Vanduijnen PT, Thole BT, Hol WGJ. Role of the Active-site helix in papain, an ab initio molecular-orbital study. Biophys Chem 1979;9(3):273–280.
63. Roux B, MacKinnon R. The cavity and pore helices the KcsA K^+ channel: electrostatic stabilization of monovalent cations. Science 1999;285:100–102.
64. Roux B, Berneche S, Im W. Ion channels, permeation and electrostatics: insight into the function of KcsA. Biochemistry 2000;39(44): 13295–13306.
65. Aqvist J, Luecke H, Quiocho FA, Warshel A. Dipoles localized at helix termini of proteins stabilize charges. Proc Natl Acad Sci 1991;88(5):2026–2030.
66. Sali D, Bycroft M, Fersht AR. Stabilization of protein structure by interaction of alpha-helix dipole with a charged side chain. Nature 1988;335(6192):740–743.

67. Quiocho FA, Sack JS, Vyas NK. Stabilization of charges on isolated ionic groups sequestered in proteins by polarized peptide units. Nature 1987;329(6139):561–564.
68. Burykin A, Warshel A. Membranes assembled from narrow carbon nanotubes block proton transport and can form effective nano filtration devices. J Comp Theo Nano 2006;3:237–242.
69. Warshel A. Calculations of enzymic reactions: calculations of pK_a, proton transfer reactions, and general acid catalysis reactions in enzymes. Biochemistry 1981;20:3167–3177.
70. Fitch C, Karp D, Gittis A, Lattman E, Garcia-Moreno B. Experimental pKa values of buried residues: analysis with continuum methods and role of water penetration. Biophys J 2002;82:3289–3304.
71. Damjanovic A, Garcia-Moreno B, Lattman EE, Garcia AE. Molecular dynamics study of water penetration in staphylococcal nuclease. Proteins 2005;60(3):433–449.
72. Kato M, Warshel A. Using a charging coordinate in studies of ionization induced partial unfolding. J Phys Chem B (in press).
73. Garcia-Moreno B, Dwyer JJ, Gittis AG, Lattman EE, Spencer DS, Stites WE. Experimental measurement of the effective dielectric in the hydrophobic core of a protein. Biophys Chem 1997;64:211–224.
74. Lodi PJ, Knowles JR. Direct evidence for the exploitation of an alpha-helix in the catalytic mechanism of triosephosphate isomerase. Biochemistry 1993;32(16):4338–4343.
75. Braun-Sand S, Warshel A. Electrostatics of proteins: principles, models and applications. In: Buchner J, Kiefhaber T, editors. Protein folding handbook. Part I. Weinheim: Wiley-VCH; 2005. p 163–200.
76. Chatelain FC, Alagem N, Xu Q, Pancaroglu R, Reuveny E, Minor DL. The pore helix dipole has a minor role in inward rectifier channel function. Neuron 2005;47(6):833–843.
77. Stanton R, Peräkylä M, Bakowies D, Kollman PA. Combined ab initio and free energy calculations to study reactions in enzymes and solution: amide hydrolysis in trypsin and aqueous solution. J Am Chem Soc 1998;120:3448–3457.
78. Shurki A, Warshel A. Structure/function correlations of proteins using MM, QM/MM and related approaches: methods, concepts, pitfalls and current progress. Adv Protein Chem 2003;66:249–312.
79. Warshel A. Energetics of enzyme catalysis. Proc Natl Acad Sci USA 1978;75:5250–5254.
80. Hwang J-K, Warshel A. Why ion pair reversal by protein engineering is unlikely to succeed. Nature 1988;334:270.