

# Amino acid deprotonation rates from classical force fields

Cite as: *J. Chem. Phys.* **157**, 085101 (2022); doi: 10.1063/5.0101960

Submitted: 5 June 2022 • Accepted: 1 August 2022 •

Published Online: 23 August 2022



View Online



Export Citation



CrossMark

Themis Lazaridis<sup>a)</sup>  and Aliasghar Sepehri

## AFFILIATIONS

Department of Chemistry, City College of New York/CUNY, 160 Convent Ave., New York, New York 10031, USA

<sup>a)</sup>Also at: Graduate Programs in Chemistry, Biochemistry, and Physics, The Graduate Center, City University of New York, 365 Fifth Ave., New York, NY 10016, USA.

Author to whom correspondence should be addressed: [tlazaridis@ccny.cuny.edu](mailto:tlazaridis@ccny.cuny.edu). Tel.: (212) 650-8364

## ABSTRACT

Acid ionization constants ( $pK_a$ 's) of titratable amino acid side chains have received a large amount of experimental and theoretical attention. In many situations, however, the rates of protonation and deprotonation,  $k_{on}$  and  $k_{off}$ , may also be important, for example, in understanding the mechanism of action of proton channels or membrane proteins that couple proton transport to other processes. Protonation and deprotonation involve the making and breaking of covalent bonds, which cannot be studied by classical force fields. However, environment effects on the rates should be captured by such methods. Here, we present an approach for estimating deprotonation rates based on Warshel's extension of Marcus's theory of electron transfer, with input from molecular simulations. The missing bond dissociation energy is represented by a constant term determined by fitting the  $pK_a$  value in solution. The statistics of the energy gap between protonated and deprotonated states is used to compute free energy curves of the two states and, thus, free energy barriers, from which the rate can be deduced. The method is applied to Glu, Asp, and His in bulk solution and select membrane proteins: the M2 proton channel, bacteriorhodopsin, and cytochrome c oxidase.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0101960>

## INTRODUCTION

Titratable residues in biomolecules play important roles in many biological processes such as enzyme catalysis, bioenergetics, and viral entry. All pH effects are mediated by such residues. In bioenergetics, they play key roles in membrane proteins that generate or consume the electrochemical proton gradient. A key property of a titratable residue is its  $pK_a$ , the pH value where the residue is protonated half of the time. The  $pK_a$  is a thermodynamic property that depends on the position of the residue and its interactions with the surroundings. Knowing the  $pK_a$  of a given residue in a given protein is valuable, but sometimes insufficient. One would also need to know the rates of protonation ( $k_{on}$ ) and deprotonation ( $k_{off}$ ) that are components of the equilibrium constant ( $K_a = k_{off}/k_{on}$ ). For example, the mechanism of action of proton pumps or proton channels involves protonation and deprotonation of key amino acids. Knowing the kinetics of these processes is key to identifying rate limiting steps or mechanisms of preventing proton backflow.

Rates of protonation and deprotonation were measured decades ago for small titratable molecules in an aqueous solution

using relaxation methods and nuclear magnetic resonance (NMR).  $k_{on}$  and  $k_{off}$  were measured as  $4.5 \cdot 10^{10} \text{ M}^{-1} \text{ s}^{-1}$  and  $7.8 \cdot 10^5 \text{ s}$ , respectively, for acetic acid, and  $1.5 \cdot 10^{10}$  and  $1.5 \cdot 10^3$ , respectively, for imidazole.<sup>1,2</sup> Thus, protonation is diffusion-limited, i.e., essentially instantaneous upon encounter, but deprotonation takes about  $1 \mu\text{s}$  for acetic acid and  $1 \text{ ms}$  for imidazole. In proteins, although  $pK_a$  measurements are abundant, kinetic data like the above are scarce. In addition, for buried residues in proteins, protonation and deprotonation require the movement of a proton into and out of the protein interior, which adds kinetic barriers. Thus, for such residues, the rate of protonation and deprotonation,  $k_{prot}$  and  $k_{deprot}$ , may differ from the overall  $k_{on}$  and  $k_{off}$ .

Many methods exist for the prediction of  $pK_a$ 's in proteins but, to our knowledge, very little theoretical work has been done on (de)protonation kinetics. The standard approach is to perform quantum mechanical calculations to obtain a free energy profile of proton dissociation, from which the rate can be inferred.<sup>3-7</sup> It is an expensive endeavor that cannot easily be applied to every situation required. Here, we explore the ability of classical force fields and simulations to estimate deprotonation rates. The method builds on the

Marcus theory of electron transfer, as extended to proton transfer and generalized for input from molecular simulations.<sup>5,8–11</sup> We apply it to amino acids in solution and to three membrane proteins.

## THEORY

Marcus's theory of electron transfer reactions treats the solvent as a dielectric continuum that responds linearly to perturbations. It predicts parabolic free energy curves of equal curvature for the reactant and product state.<sup>12</sup> The barriers of the reaction are predicted to be

$$\Delta G^\ddagger = (\lambda \pm \varepsilon)^2 / 4\lambda, \quad (1)$$

where  $\varepsilon = \Delta G$ , the free energy change of the reaction, and  $\lambda$  is the “reorganization energy.” These two parameters are related to the energy gaps  $\Delta_1$  and  $\Delta_2$  between the two states at the bottom of the two wells (Fig. 1):

$$\Delta_1 = \lambda + \varepsilon \quad \Delta_2 = \lambda - \varepsilon. \quad (2)$$

Thus, calculating  $\Delta_1$  and  $\Delta_2$  as the most frequent values observed upon simulating the two states, the above equations yield  $\lambda$ ,  $\varepsilon$ , and the reaction barriers. This is what we refer to as “simple Marcus model.”

Much work has been done on testing and extending Marcus's theory using molecular simulations.<sup>8–10,13–15</sup> After Warshel's suggestion,<sup>9</sup> the energy gap between the two states has become the most commonly used reaction coordinate. The assumption of linear response, which leads to parabolic profiles of equal curvature, has been relaxed,<sup>16,17</sup> and the temperature dependence of  $\lambda$  has been considered.<sup>18</sup> Although proton and electron transfer differ in important ways, they have in common the key role of solvent reorganization. This allowed the application of the same approach to proton or H atom transfer reactions.<sup>5,8,19,20</sup>

One difficulty that this method commonly encounters is that the small values of  $\Delta$  may not be sampled by the simulation of either

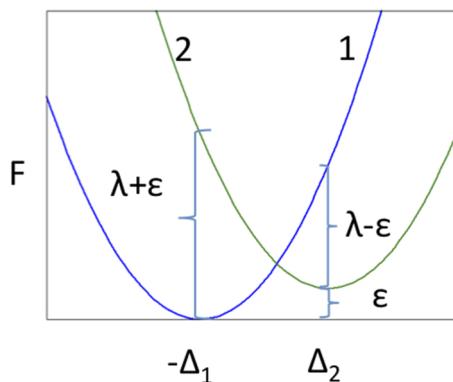


FIG. 1. Schematic of Marcus parabolas for the free energy of two states 1 and 2 and definition of reorganization energy  $\lambda$  and free energy change  $\varepsilon$ .  $\Delta_1$  is defined as  $F_2 - F_1$  with state 1 sampled and  $\Delta_2$  as  $F_1 - F_2$  with state 2 sampled. The x-axis is  $F_1 - F_2$ .

state. One can either fit and extrapolate the curves or calculate the curves at small  $\Delta$  values by performing simulations of a hybrid system and using umbrella sampling to obtain the full curve.<sup>5,8,11,13,21,22</sup> Both approaches are evaluated here.

## METHODS

The energy gap between protonated and deprotonated states was calculated with the MOBHY module<sup>23</sup> implemented in a custom version of the program Chemistry at Harvard Molecular Mechanics (CHARMM), version c41a1.<sup>24</sup> For deprotonation, the protonated amino acid was simulated and periodic attempts were made to move the proton to a neighboring, H-bonded water molecule in ideal  $\text{H}_3\text{O}^+$  geometry (the closest of two possible positions is chosen). For protonation, the simulation was carried out with deprotonated amino acid H-bonded to hydronium. Upon protonation, the H is placed according to the standard internal coordinates of the side chain. An H-bond is considered to exist when the donor–acceptor (DA) distance is less than 3.1 (Glu, Asp) or 3.2 (His) Å, and the D-H-A angle is at least 90°. No hops were accepted, only the hypothetical change in energy upon a proton hop attempt was calculated,  $\Delta_d$  for deprotonation and  $\Delta_p$  for protonation. The excess proton was modeled as a classical hydronium,<sup>25</sup> and the CHARMM version of the TIP3 model was used for water. Because the classical force field energies ( $\Delta E_{\text{mm}}$ ) do not include the energy of covalent bonds being broken, the energy of the deprotonated state has to be corrected by a quantity that reflects the energy of breaking the bond of H with the side chain and of forming a bond with water,  $\Delta E = \Delta E_{\text{mm}} + E_{\text{b,sch}} - E_{\text{b,H}_3\text{O}}$ . The value 165 kcal/mol was used for  $E_{\text{b,H}_3\text{O}}$  and  $E_{\text{b,sch}}$  was fit to solvation free energy calculations of the relevant amino acid in an aqueous solution.<sup>23</sup> The values used here, derived for use with the CHARMM 36 force field, are 316 kcal/mol for Glu, 303 for Asp, 253 for Hsd ( $\pi$  tautomer, proton on Nδ), and 272 for Hse ( $\tau$  tautomer, proton on Ne).

We first considered blocked amino acids in solution. A Glu, Asp, or His amino acid acetylated at the N-terminus and methylated at the C-terminus was placed in an 18.856-Å cubic box with 217 TIP3 molecules (and 1  $\text{Cl}^-$  in the case of His). The  $\Delta$  distributions were calculated by attempting proton hops every ten steps in a 1-ns constant pressure simulation at 300 K with Particle Mesh Ewald (PME) and a 2-fs timestep. In the case of His, because the  $\text{H}_3\text{O}^+$  tended to drift off, a distance restraint was used to keep it close to the His. Separate calculations were done for the two His tautomers. To obtain the distribution of  $\Delta$  for small values of  $\Delta$ , a hybrid system was also simulated using the PERT module of CHARMM. The free energy of the protonated state is  $-kT \ln p(\Delta)$ , that of the deprotonated state  $-kT \ln p(\Delta) + c_1$ , with  $c_1$  chosen so that the curve matches the protonated state curve at  $\Delta = 0$ , and that of the hybrid state  $-kT \ln p(\Delta) - \xi \cdot \Delta_1 + c_2$ ,<sup>22</sup> where  $\xi$  is the coupling parameter (0–1) and  $c_2$  is chosen to match the protonated state curve in the region where they overlap.

We then examined three widely studied membrane proteins: the influenza M2 proton channel, bacteriorhodopsin, and cytochrome c oxidase. For the M2 channel, the 6BKK crystal structure<sup>26</sup> was embedded in a lipid bilayer of 138 palmitoyl-oleoyl-phosphatidylcholine (POPC) lipids and about 4060 TIP3 water molecules in a  $70 \times 70 \times 63$  Å<sup>3</sup> box. For bacteriorhodopsin, we used the crystal structure of the D85S mutant (1JV7<sup>27</sup>), considered to

be a model for the O state, with S85 reverted to Asp. The missing loop 64–77 was built from pdb 1QHJ. The structure was placed in a lipid bilayer with 114 POPC lipids, crystal waters, 6500 additional waters, and 1  $K^+$ . The energy gap was calculated every ten steps in a 200 ps simulation. A hybrid state was also simulated to obtain data for low energy gap values. That state consisted of 50% protonated Asp85 H-bonded to water and 50% deprotonated Asp85 H-bonded to  $H_3O^+$ .

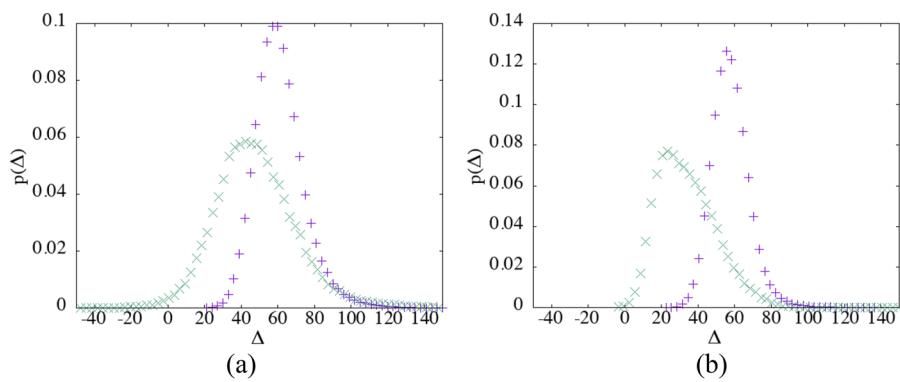
For cytochrome c oxidase, we used the crystal structure 2GSM,<sup>28</sup> which includes the essential subunits I and II. It was equilibrated in a dimyristoylphosphatidylcholine (DMPC) membrane with 144 lipids and about 16 300 TIP3 water molecules. Parameters for the cofactors were those developed by Johansson *et al.*<sup>29</sup> The first state modeled was Pm (heme an oxidized, CuA oxidized, heme a3 oxoferryl with ionic propionates, and CuB oxidized with  $OH^-$ ). Calculations were subsequently done for reduced heme a or reduced heme a3. Deprotonation of Glu 286 was considered in the crystal structure (with the Glu side chain pointing down into the D channel) or in an alternative conformation with Glu

pointing up and four extra waters added in the hydrophobic cavity that lies above Glu. The upward conformation was generated by rotating the  $\chi_2$  dihedral angle to  $180^\circ$  and is similar to states considered in previous work.<sup>30</sup> Hybrid states were also simulated as above. In all cases, we used the CHARMM 36 force field,<sup>31</sup> a 12-Å cutoff for van der Waals interactions, and PME for long range electrostatics. The membrane system construction and equilibration were done using the CHARMM-gui server.<sup>32</sup>

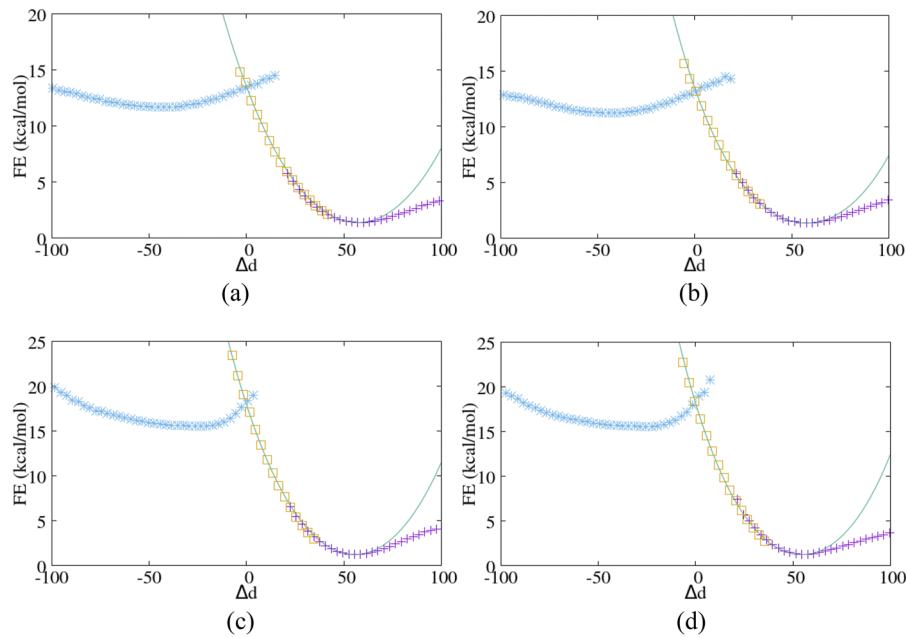
## RESULTS

### Deprotonation in aqueous solution

The distributions of energy gap  $p(\Delta)$  for protonation and deprotonation of Glu and Hse ( $\tau$ ) in solution are shown in Fig. 2 [plots for Asp and Hsd ( $\pi$ ) are very similar and are not shown]. Deviations from Gaussian behavior can be seen, especially for protonation, which is reflected in that the corresponding free energy curves in Fig. 3 are not perfectly parabolic.



**FIG. 2.** Probability distribution of the energy gap for (a) Glu deprotonation (purple crosses) and protonation (green Xs), (b) His<sup>+</sup> deprotonation to produce Hse (purple crosses) and Hse protonation (green Xs).  $\Delta$  is the energy gap to the other state ( $E_{\text{deprotonated}} - E_{\text{protonated}}$  for deprotonation and  $E_{\text{protonated}} - E_{\text{deprotonated}}$  for protonation).



**FIG. 3.** Free energy curves for (a) Glu, (b) Asp, (c) Hse, and (d) Hsd as a function of the energy gap  $\Delta_d = E_{\text{deprotonated}} - E_{\text{protonated}}$ . Blue X denotes the protonated state curve, purple crosses and orange squares are the deprotonated state curve (orange squares are the hybrid state data), and green continuous lines are quartic fits of the inner branch of the protonated state curve.

**TABLE I.** Free energies ( $\Delta G$ ) and barriers ( $\Delta G^\ddagger$ ) from different methods (kcal/mol).

	FE curves			Simple Marcus model				Experiment <sup>a</sup>
	$\Delta G$	$\Delta G^\ddagger_p$	$\Delta G^\ddagger_d$	$\Delta G$	$\lambda$	$\Delta G^\ddagger_p$	$\Delta G^\ddagger_d$	$\Delta G^\ddagger_d$
Glu	10.3	1.5	11.8	7.2	50.0	9.0	16.2	9
Asp	9.9	1.9	11.8	8.0	49.5	8.7	16.7	9
Hse	14.3	2.6	16.9	16.0	39.5	3.5	19.5	14
Hsd	14.3	2.5	16.8	14.5	40.0	4.1	18.6	14

<sup>a</sup>For acetic acid and imidazole.<sup>1,2</sup>

The protonation curves reach  $\Delta = 0$ , but the deprotonation curves do not. To complete the curves for low values of  $\Delta$ , we simulated a hybrid state at either  $\xi = 0.5$  for His or  $\xi = 0.8$  for Glu & Asp ( $\xi = 1$  corresponds to the protonated state) and obtained additional data that allowed us to construct the full free energy curves, which are obtained by matching the protonation and deprotonation diabatic branches at  $\Delta = 0$  (Fig. 3). The protonation curves (blue X) are shallow. The deprotonation curves have much “softer” outer branches (large  $\Delta$ , purple +), while the inner branches are steeper and closer to parabolic. Parabolic fitting of the low- $\Delta$  branch of the  $FE(\Delta)$  curve and extrapolation to  $\Delta = 0$  is able to reproduce the hybrid state umbrella sampling results quite well, but including a quartic term improves the fit.

The complete curves are used to predict the  $\Delta G$ s and barriers shown in Table I. For comparison, if we use the experimental deprotonation rates for acetic acid and imidazole (see Introduction) as representative of Asp/Glu and His, and transition state theory [ $k = k_B T/h \exp(-\Delta G^\ddagger/kT)$ ], we obtain a deprotonation barrier of 9 kcal/mol for Asp/Glu and 14 kcal/mol for His. The deprotonation barriers that we obtain from the classical force field and the present method are about 3 kcal/mol higher than that. Estimates of the barriers from the simple Marcus model [Eqs. (1) and (2)] are even further exaggerated (Table I).

### Deprotonation in the gas phase

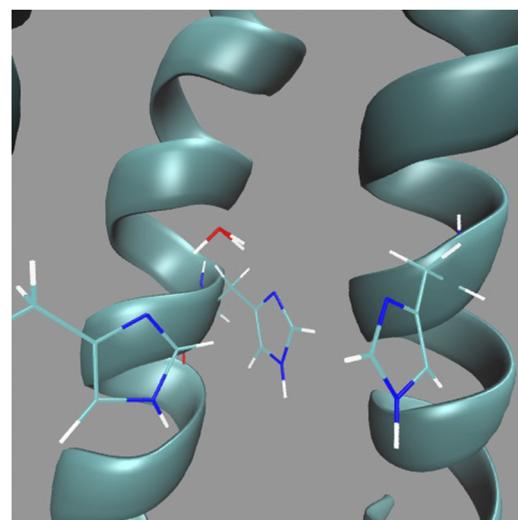
Even if classical force fields cannot reproduce the absolute rates of deprotonation, they should be able to reproduce the relative change in rates as one moves from aqueous solutions to complex biological environments, in much the same way that classical energy functions and continuum solvation can reproduce shifts in  $pK_a$  values relative to solution.<sup>33,34</sup> Here, we examine the effect of environment on the calculated barriers, starting from amino acids in the gas phase.

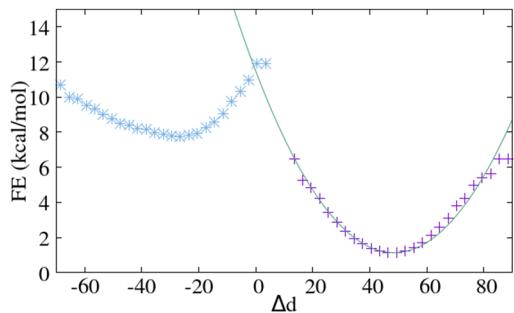
For Glu in the gas phase with one water molecule H-bonded to it, the distribution of  $\Delta$  is shifted to much larger values. For deprotonation, we obtain a  $\langle \Delta_d \rangle$  of 112 kcal/mol, ranging from 84 to 148. For protonation, we obtain a  $\langle \Delta_p \rangle$  of -49 kcal/mol, ranging from -127 to +50. By extrapolating quadratically the free energy curve (not shown), we estimate a deprotonation barrier of 49 kcal/mol and a deprotonated state 47 kcal/mol higher than the protonated one. It is clear that the deprotonated state is highly destabilized by the lack of surrounding water molecules. Similar results are obtained for Hsd but with a narrower range of  $\Delta$  values:  $\langle \Delta_d \rangle = 73$  (from 54 to 88) and  $\langle \Delta_p \rangle = -32.5$  (from -55 to -16) kcal/mol. Extrapolation gives

a deprotonation barrier of about 64 kcal/mol, a substantial protonation barrier of 22 kcal/mol, and a difference of 42 kcal/mol in free energy. Thus, the classical treatment can qualitatively reproduce the effect of the gas-phase environment.

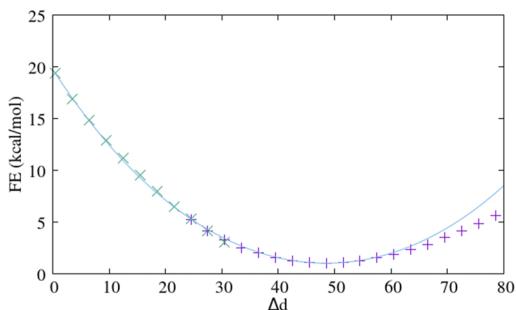
### His37 titration in the influenza M2 proton channel

The M2 proton channel of influenza<sup>35</sup> is a membrane-embedded helical tetramer with a quartet of His residues in its interior that are responsible for its low-pH activation and its proton selectivity.<sup>36,37</sup> The  $pK_a$ ’s of these His have been measured experimentally,<sup>38–41</sup> as well as by theoretical calculations,<sup>42,43</sup> yielding a variety of results. Deprotonation rates have been calculated using transition state theory and Quantum Mechanics/Molecular Mechanics (QM/MM) energy profiles.<sup>44</sup> Here, we apply our approach on the first titration of the His tetrad (protonating a fully neutral His ring). For protonation, the four His are in the Hse tautomer and are approached by a  $H_3O^+$  in the outer vestibule, which can protonate the unprotonated N $\delta$ . For deprotonation, one of the His is doubly protonated and can deprotonate either the N $\delta$  or the N $\epsilon$  (Fig. 4).

**FIG. 4.** The His 37 residues in the M2 proton channel (the fourth helix is omitted for clarity).



**FIG. 5.** FE curves for protonation and deprotonation of N $\delta$  in M2. The green line  $[1.13 + 0.0044(\Delta - 48.5)^2]$  is a fit and extrapolation of the inner branch of the deprotonation curve.



**FIG. 7.** Deprotonation free energy of D85 as a function of the energy gap in bacteriorhodopsin. Data marked X are from the hybrid simulation. The line is  $1.04 + 0.00725(\Delta - 48.5)^2 + 3 \cdot 10^{-7}(\Delta - 48.5)^4$ .

The free energy curves for protonation and deprotonation of N $\delta$  are shown in Fig. 5. The deprotonation curve of N $\epsilon$  is similar. The inner branch of the curve reaches down to 13 kcal/mol and is fitted well by a parabolic line. Extrapolation to  $\Delta = 0$  yields barriers of 4.0 and 10.3 kcal/mol for N $\delta$  protonation and deprotonation, respectively. The barrier for N $\epsilon$  deprotonation is 10.6 kcal/mol. The  $\Delta_{\min}$  (position of the free energy minimum) for deprotonation is lower than in bulk water and that for protonation is higher. The resulting barriers are lower than in bulk water for deprotonation and higher for protonation. The first shift is apparently due to the lower solvent reorganization cost in the channel interior, and the second shift is due to the lack of hydration on the other side of the neutral His layer, which lowers the stability of the charged form. The barriers for deprotonation from N $\delta$  to N $\epsilon$  are basically the same.

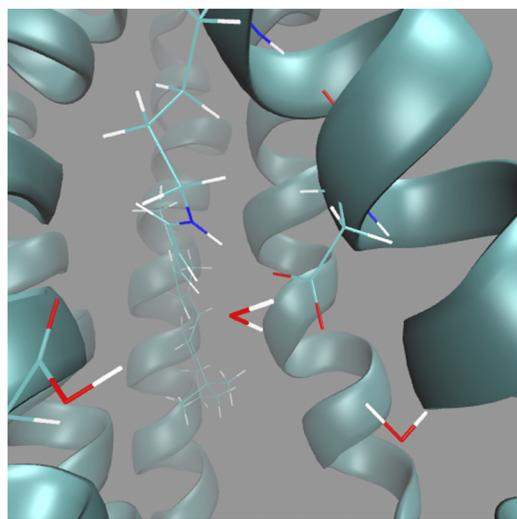
We should note that the rates we would obtain here applying transition state theory do not correspond to the  $k_{\text{off}}$  and  $k_{\text{on}}$  that combine to give  $K_a$  but to the first or last step,  $k_{\text{deprot}}$ ,  $k_{\text{prot}}$ . What

we calculate is the rate of deprotonation of one His, whereas  $k_{\text{off}}$  is the rate of release of the proton into the bulk. To reach the bulk, the proton needs to surmount additional barriers. We also need to take into account the “degeneracy” generated by the presence of four His. Similarly,  $K_{\text{on}}$  is the rate of protonation from the bulk and includes additional barriers (diffusion and other channel constrictions). What we calculate is closer to the rate of proton exchange determined by NMR.<sup>40</sup> That rate was consistent with a barrier of about 10 kcal/mol.

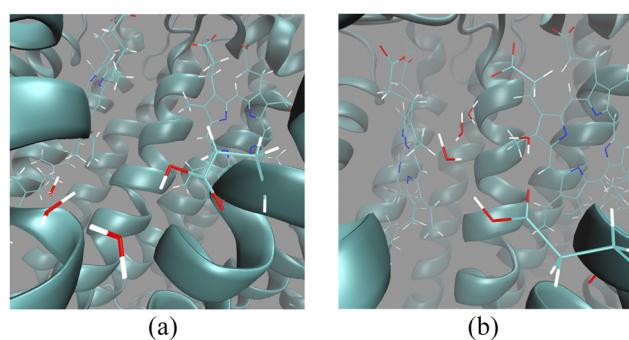
### Asp85 deprotonation in bacteriorhodopsin

The last step of the photocycle of the light-driven proton pump bacteriorhodopsin ( $O \rightarrow BR$  transition) involves the transfer of a proton from Asp85 to the proton release group, which is likely a pair of Glu and the neighboring waters.<sup>45</sup> An alternative mechanism involves abstraction of the proton by an OH<sup>-</sup> originating from the extracellular side.<sup>46-48</sup> Here, we consider the deprotonation of Asp85 in the O state where the Lys216-retinal Schiff base has been reprotominated and is in a trans form. The nearby Asp212 and the Glu of the proton release group are deprotonated. We use the crystal structure of the D85S mutant (pdb 1JV7,<sup>27</sup>), which is thought to be a model for the O state, after reverting S85 to D (Fig. 6).

The deprotonation free energy curve is shown in Fig. 7 and predicts a barrier of 19 kcal/mol. The time associated with such a



**FIG. 6.** Key residues in bacteriorhodopsin. Asp85 on the left, Asp212 on the right, and Lys216-retinal in the background.

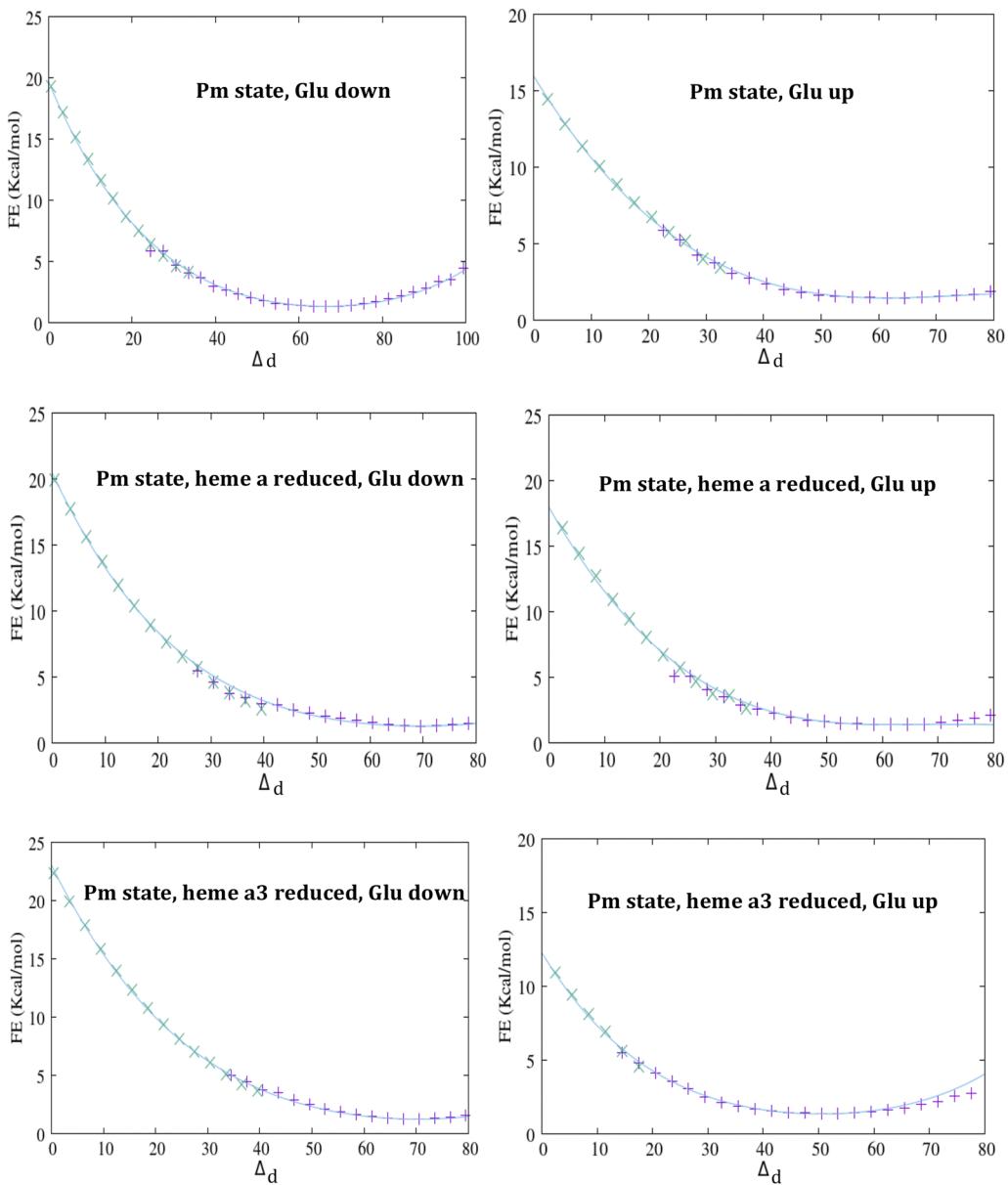


**FIG. 8.** Glu286 in CCO in two configurations, (a) crystal (“down”), and (b) “up.” In the background are heme a (left) and heme a3 (right).

barrier is 9 s. If, in view of the overestimation of the barriers in bulk water, we consider the barrier to be lower by 3 kcal/mol, we obtain 61 ms, which is somewhat longer than the 10 ms that the overall O → BR transition takes,<sup>49</sup> which includes additional difficult steps for the proton, such as crossing the Arg82 region. This indicates that the OH<sup>-</sup> abstraction mechanism may be more plausible kinetically. This will have to be verified by separate calculations. This is one example where calculating deprotonation rates may allow us to select one mechanism over another.

### Glu286 deprotonation in cytochrome c oxidase

Cytochrome c oxidase is the terminal enzyme of the respiratory chain that contributes majorly to generating the electrochemical gradient that fuels ATP synthesis.<sup>50</sup> Electrons from cytochrome c are transferred to a binuclear Fe–Cu center via two Cu and two hemes where they are used to reduce O<sub>2</sub> to water using four protons from the cell or mitochondrial interior. Another four protons are transferred from the interior to the external side of the membrane. Key



**FIG. 9.** Deprotonation free energy curves for E286 in cytochrome c oxidase in different states. Green Xs are from the  $\xi = 0.5$  hybrid state simulation. The curves include a quadratic and a quartic term, except for Pm "up" and reduced heme an "up" where a cubic term gives a better fit than the quartic.

**TABLE II.** Mean energy gaps and deprotonation barriers (kcal/mol) of E286 in cytochrome c oxidase in the Pm state and in the Pm state after reduction of heme a or heme a3.

	Down		Up	
	$\langle \Delta \rangle$	Barrier	$\langle \Delta \rangle$	Barrier
Pm	65.7	18.3	63.0	14.5
Pm red a	68.5	19.2	60.9	16.6
Pm red a3	69.6	21.8	52.0	10.9

to the function of this enzyme is a central Glu residue (here E286) that appears to be a transition hub for the protons. Here, we calculate the deprotonation rate of this Glu in two conformations: “down” (as in the crystal structure, pointing inward toward the D channel) and “up,” with the Glu pointing outward with four water molecules added in the hydrophobic cavity (Fig. 8). This movement of Glu has been hypothesized to provide a “valve” preventing the backflow of protons.<sup>30</sup> We do this calculation first in the Pm state (see Methods) and then investigate the effect of electron transfer on deprotonation rates by reducing either heme a or heme a3. Because the low  $\Delta$  values are not sampled well, we also simulate a hybrid state at  $\xi = 0.5$  with PERT.

The deprotonation free energy curves are shown in Fig. 9 and the barriers are shown in Table II. In the Pm state, the energy gap statistics are somewhat more favorable in the up orientation. Reduction of heme a increases both barriers slightly. Reduction of heme a3 has a larger effect, increasing the downward barrier and lowering the upward barrier. The reason for the change in barriers upon reduction is simple electrostatics: placing a negative charge in the direction of  $H^+$  movement favors deprotonation; placing it in the opposite direction disfavors deprotonation. In most cases, the mean energy gap correlates with the barrier except for the up configuration in reduced heme a where the  $\langle \Delta \rangle$  decreases but the barrier increases due to increased steepness of the curve.

## DISCUSSION

In this paper, we adapted a method to obtain free energy curves, barriers, and rates for electron and proton transfer in simple systems to titration of amino acids in proteins. Although classical force fields are not expected to provide accurate absolute rates, they could account for the effect of the environment, such as long-range interactions, on the rates. The method does not require lengthy simulations. In some cases, an additional hybrid state simulation may be required to obtain the full free energy curve, but such tools are available in most biomolecular simulation packages. This type of calculations could provide insights into systems where proton transfer is central, such as proton pumps and proton channels. The results could be useful in their own right, and they could also provide rate data for hyperdynamics-style MD simulations.<sup>51</sup> The main hindrance in this line of work is the lack of experimental data to validate predictions. Hopefully, the availability of theoretical estimates would stimulate experimental labs to take up this important task.

The emphasis in this paper has been on deprotonation because protonation is usually very fast. In fact, in solution, it is close to

diffusion-limited. So, to compute  $k_{on}$  one has to account for the diffusional encounter rate.<sup>52–54</sup> In solution, deprotonation is likely to be the rate-limiting step, so  $k_{deprot}$  should be a good approximation to  $k_{off}$ . For sites buried in proteins, the situation is much more complex. A proton released to or absorbed from the bulk solvent has to traverse significant barriers. In addition, the multiplicity of paths through the anisotropic protein matrix has to be accounted for, and each path must be weighted according to each probability.

The results on cytochrome c oxidase allow us to make an interesting point. The  $pK_a$  is a valuable thermodynamic quantity that reflects the free energy of removing a proton from a given site and placing it in the bulk. The approach of a negative charge raises the  $pK_a$  of an acidic site. However, the effect on  $k_{off}$  and  $k_{on}$  is not isotropic.  $K_{off}$  actually increases in the direction of the negative charge and decreases in the opposite direction. This could be important in tracing proton paths and estimating their kinetic feasibility.

## ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (Grant No. MCB-1855942). We thank Dr. M. Gunner and Dr. Q. Cui for CHARMM topology and parameters for the cytochrome c oxidase cofactors.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Themis Lazaridis:** Conceptualization (lead); Writing – original draft (lead); Writing – review & editing (lead). **Aliasghar Sepehri:** Data curation (supporting); Investigation (supporting).

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

1. M. Eigen, W. Kruse, G. Maass, and L. De Maeyer, *Progress in Reaction Kinetics*, edited by G. Porter (Macmillian, New York, 1964), Vol. 2, pp. 285–318.
2. E. K. Ralph and E. Grunwald, *J. Am. Chem. Soc.* **91**, 2422 (1969).
3. J. M. Park, A. Laio, M. Iannuzzi, and M. Parrinello, *J. Am. Chem. Soc.* **128**, 11318 (2006).
4. R. Liang, J. M. J. Swanson, Y. Peng, M. Wikström, and G. A. Voth, *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7420 (2016).
5. K. Ando and J. T. Hynes, *J. Phys. Chem. B* **101**, 10464 (1997).
6. K. Ando and J. T. Hynes, *J. Phys. Chem. A* **103**, 10398 (1999).
7. P. Goyal, S. Yang, and Q. Cui, *Chem. Sci.* **6**, 826 (2015).
8. A. Warshel, *J. Phys. Chem.* **86**, 2218 (1982).
9. J. K. Hwang and A. Warshel, *J. Am. Chem. Soc.* **109**, 715 (1987).
10. E. A. Carter and J. T. Hynes, *J. Phys. Chem.* **93**, 2184 (1989).
11. G. King and A. Warshel, *J. Chem. Phys.* **93**, 8682 (1990).
12. R. A. Marcus, *Annu. Rev. Phys. Chem.* **15**, 155 (1964).

- <sup>13</sup>R. A. Kuharski, J. S. Bader, D. Chandler, M. Sprik, M. L. Klein, and R. W. Impey, *J. Chem. Phys.* **89**, 3248 (1988).
- <sup>14</sup>W. W. Parson, *J. Chem. Phys.* **152**, 184106 (2020).
- <sup>15</sup>J. B. Straus and G. A. Voth, *J. Phys. Chem.* **97**, 7388 (1993).
- <sup>16</sup>H. X. Zhou and A. Szabo, *J. Chem. Phys.* **103**, 3481 (1995).
- <sup>17</sup>T. Ichiye, *J. Chem. Phys.* **104**, 7561 (1996).
- <sup>18</sup>P. K. Ghorai and D. V. Matyushov, *J. Phys. Chem. A* **110**, 8857 (2006).
- <sup>19</sup>R. A. Marcus, *Faraday Symp. Chem. Soc.* **10**, 60 (1975).
- <sup>20</sup>J. M. Mayer, *Acc. Chem. Res.* **44**, 36 (2011).
- <sup>21</sup>G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
- <sup>22</sup>K. Ando and S. Kato, *J. Chem. Phys.* **95**, 5966 (1991).
- <sup>23</sup>T. Lazaridis and G. Hummer, *J. Chem. Inf. Model.* **57**, 2833 (2017).
- <sup>24</sup>B. R. Brooks, C. L. Brooks III, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, *J. Comput. Chem.* **30**, 1545 (2009).
- <sup>25</sup>D. E. Sagnella and G. A. Voth, *Biophys. J.* **70**, 2043 (1996).
- <sup>26</sup>J. L. Thomaston, N. F. Polizzi, A. Konstantinidi, J. Wang, A. Kolocuris, and W. F. DeGrado, *J. Am. Chem. Soc.* **140**, 15219 (2018).
- <sup>27</sup>S. Rouhani, J.-P. Cartailler, M. T. Facciotti, P. Walian, R. Needleman, J. K. Lanyi, R. M. Glaeser, and H. Luecke, *J. Mol. Biol.* **313**, 615 (2001).
- <sup>28</sup>L. Qin, C. Hiser, A. Mulichak, R. M. Garavito, and S. Ferguson-Miller, *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16117 (2006).
- <sup>29</sup>M. P. Johansson, V. R. I. Kaila, and L. Laakkonen, *J. Comput. Chem.* **29**, 753 (2008).
- <sup>30</sup>V. R. I. Kaila, M. I. Verkhovsky, G. Hummer, and M. Wikström, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6255 (2008).
- <sup>31</sup>R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, *J. Chem. Theory Comput.* **8**, 3257 (2012).
- <sup>32</sup>E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda, and W. Im, *J. Comput. Chem.* **35**, 1997 (2014).
- <sup>33</sup>E. Alexov, E. L. Mehler, N. Baker, A. M. Baptista, Y. Huang, F. Milletti, J. Erik Nielsen, D. Farrell, T. Carstensen, M. H. M. Olsson, J. K. Shen, J. Warwicker, S. Williams, and J. M. Word, *Proteins: Struct., Funct., Bioinf.* **79**, 3260 (2011).
- <sup>34</sup>M. R. Gunner and N. A. Baker, *Continuum Electrostatics Approaches to Calculating  $pK_a$ s and  $E_m$ s in Proteins*, 1st ed. (Elsevier, 2016).
- <sup>35</sup>M. Hong and W. F. DeGrado, *Protein Sci.* **21**, 1620 (2012).
- <sup>36</sup>L. H. Pinto, L. J. Holsinger, and R. A. Lamb, *Cell* **69**, 517 (1992).
- <sup>37</sup>C. Wang, R. A. Lamb, and L. H. Pinto, *Biophys. J.* **69**, 1363 (1995).
- <sup>38</sup>J. Hu, R. Fu, K. Nishimura, L. Zhang, H.-X. Zhou, D. D. Busath, V. Vijayvergiya, and T. A. Cross, *Proc. Natl. Acad. Sci. U. S. A.* **103**, 6865 (2006).
- <sup>39</sup>Y. Miao, R. Fu, H.-X. Zhou, and T. A. Cross, *Structure* **23**, 2300 (2015).
- <sup>40</sup>F. Hu, K. Schmidt-Rohr, and M. Hong, *J. Am. Chem. Soc.* **134**, 3703 (2012).
- <sup>41</sup>M. T. Colvin, L. B. Andreas, J. J. Chou, and R. G. Griffin, *Biochemistry* **53**, 5987 (2014).
- <sup>42</sup>W. Chen, Y. Huang, and J. Shen, *J. Phys. Chem. Lett.* **7**, 3961 (2016).
- <sup>43</sup>H. Torabifard, A. Panahi, and C. L. Brooks, *Proc. Natl. Acad. Sci. U. S. A.* **117**, 3583 (2020).
- <sup>44</sup>R. Liang, J. M. J. Swanson, J. J. Madsen, M. Hong, W. F. DeGrado, and G. A. Voth, *Proc. Natl. Acad. Sci.* **112**, E6955 (2016).
- <sup>45</sup>J. K. Lanyi, *Annu. Rev. Physiol.* **66**, 665 (2004).
- <sup>46</sup>H. Luecke, *Biochim. Biophys. Acta, Bioenerg.* **1460**, 133 (2000).
- <sup>47</sup>M. T. Facciotti, S. Rouhani, and R. M. Glaeser, *FEBS Lett.* **564**, 301 (2004).
- <sup>48</sup>J. Herzfeld and J. C. Lansing, *Annu. Rev. Biophys. Biomol. Struct.* **31**, 73 (2002).
- <sup>49</sup>S. P. Balashov, M. Lu, E. S. Imasheva, R. Govindjee, T. G. Ebrey, B. Othersen, Y. Chen, R. K. Crouch, and D. R. Menick, *Biochemistry* **38**, 2026 (1999).
- <sup>50</sup>M. Wikström, K. Krab, and V. Sharma, *Chem. Rev.* **118**, 2469 (2018).
- <sup>51</sup>A. F. Voter, *Phys. Rev. Lett.* **78**, 3908 (1997).
- <sup>52</sup>B. A. Luty, S. El Amrani, and J. A. McCammon, *J. Am. Chem. Soc.* **115**, 11874 (1993).
- <sup>53</sup>B. R. Jagger, A. A. Ojha, and R. E. Amaro, *J. Chem. Theory Comput.* **16**, 5348 (2020).
- <sup>54</sup>S. K. Sadiq, A. Muñiz Chicharro, P. Friedrich, and R. C. Wade, *J. Chem. Theory Comput.* **17**, 7912 (2021).