

“Proton Holes” in Long-Range Proton Transfer Reactions in Solution and Enzymes: A Theoretical Analysis

Demian Riccardi,[†] Peter König,^{†,‡,||} Xavier Prat-Resina,[†] Haibo Yu,[†] Marcus Elstner,[‡] Thomas Frauenheim,^{‡,§} and Qiang Cui^{*,†}

Contribution from the Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Avenue, Madison, Wisconsin 53706, and Theoretische Physik, Universität Paderborn, Warburger Strasse 100, 33098 Paderborn, Germany

Received August 8, 2006; E-mail: cui@chem.wisc.edu

Abstract: Proton transfers are fundamental to chemical processes in solution and biological systems. Often, the well-known Grotthuss mechanism is *assumed* where a series of *sequential* “proton hops” initiates from the donor and combines to produce the net transfer of a positive charge over a long distance. Although direct experimental evidence for the sequential proton hopping has been obtained recently, alternative mechanisms may be possible in complex molecular systems. To understand these events, all accessible protonation states of the mediating groups should be considered. This is exemplified by transfers through water where the individual water molecules can exist in three protonation states (water, hydronium, and hydroxide); as a result, an alternative to the Grotthuss mechanism for a proton transfer through water is to generate a hydroxide by first protonating the acceptor and then transfer the hydroxide toward the donor through water. The latter mechanism can be most generally described as the transfer of a “proton hole” from the acceptor to the donor where the “hole” characterizes the deprotonated state of any mediating molecule. This pathway is distinct and is rarely considered in the discussion of proton-transfer processes. Using a calibrated quantum mechanical/molecular mechanical (QM/MM) model and an effective sampling technique, we study proton transfers in two solution systems and in Carbonic Anhydrase II. Although the relative weight of the “proton hole” and Grotthuss mechanisms in a specific system is difficult to determine precisely using any computational approach, the current study establishes an energetics motivated framework that hinges on the donor/acceptor pK_a values and electrostatics due to the environment to argue that the “proton hole” transfer is likely as important as the classical Grotthuss mechanism for proton transport in many complex molecular systems.

I. Introduction

The detailed molecular mechanism of long-range proton-transfer processes in solution¹ and biological systems² remains elusive despite the large body of studies^{2–6} motivated by the prominent role of proton transport in bioenergy transduction⁷ and technological applications.⁸ The large number of chemical groups that potentially participate in the transfer process makes it challenging to pinpoint the transfer mechanism and critical factors that limit the corresponding energetics and kinetics. The

fact that the dynamical fluctuations of these groups are likely important further complicates both experimental and theoretical analyses. Nevertheless, the classical Grotthuss mechanism⁹ is commonly assumed (Figure 1a), where sequential “proton hops”¹⁰ between the initial donor and the ultimate acceptor are mediated by water molecules and/or other ionizable groups. Indeed, rather long “water-wires”, intermittently mixed with titratable amino acid sidechains, are commonly seen in the X-ray structure of biomolecules that transport protons,^{2,3} and the spectroscopic signature of protonated water molecules has been captured in the biomolecular proton pump, bacteriorhodopsin.¹¹ Theoretical analyses in the Grotthuss framework have supported the mechanism by reproducing a large body of kinetic data related to proton transfers in the condensed phase, such as proton diffusion in bulk water⁶ and transport properties in a number of systems.^{4,6,12} Perhaps due to these successes, mechanisms that involve the deprotonated form of mediating groups, such

[†] University of Wisconsin.

[‡] Universität Paderborn.

[§] Current address: Bremen Center for Computational Material Science, Universität Bremen, 28334 Bremen, Germany.

^{||} Current address: The BACTER Institute, University of Wisconsin, Madison, Wisconsin 53706.

(1) Ando, K.; Hynes, J. T. *Adv. Chem. Phys.* **1999**, *110*, 381.

(2) Schultz, B. E.; Chan, S. I. *Ann. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 23.

(3) Luecke, H.; Lanyi, J. K. *Adv. Prot. Chem.* **2003**, *63*, 111.

(4) Roux, B. *Acc. Chem. Res.* **2002**, *35*, 366.

(5) Braun-Sand, S.; Strajbl, M.; Warshel, A. *Biophys. J.* **2004**, *87*, 2221.

(6) Voth, G. A. *Acc. Chem. Res.* **2006**, *39*, 143.

(7) Nicholls, D. G.; Ferguson, S. J. *Bioenergetics 3*; Academic Press: New York, 2002.

(8) Kreuer, K.; Paddison, S. J.; Spohr, E.; Schuster, M. *Chem. Rev.* **2004**, *104*, 4637.

(9) de Grotthuss, C. J. T. *Ann. Chim.* **1806**, *58*, 54.

(10) Mohammed, O. F.; Pines, D.; Dreyer, J.; Pines, E.; Nibbering, E. T. J. *Science* **2005**, *310*, 83.

(11) Garczarek, F.; Gerwert, K. *Nature* **2006**, *439*, 109.

(12) Sham, Y. Y.; Muegge, I.; Warshel, A. *Proteins: Struct., Funct., Genet.* **1999**, *36*, 484.

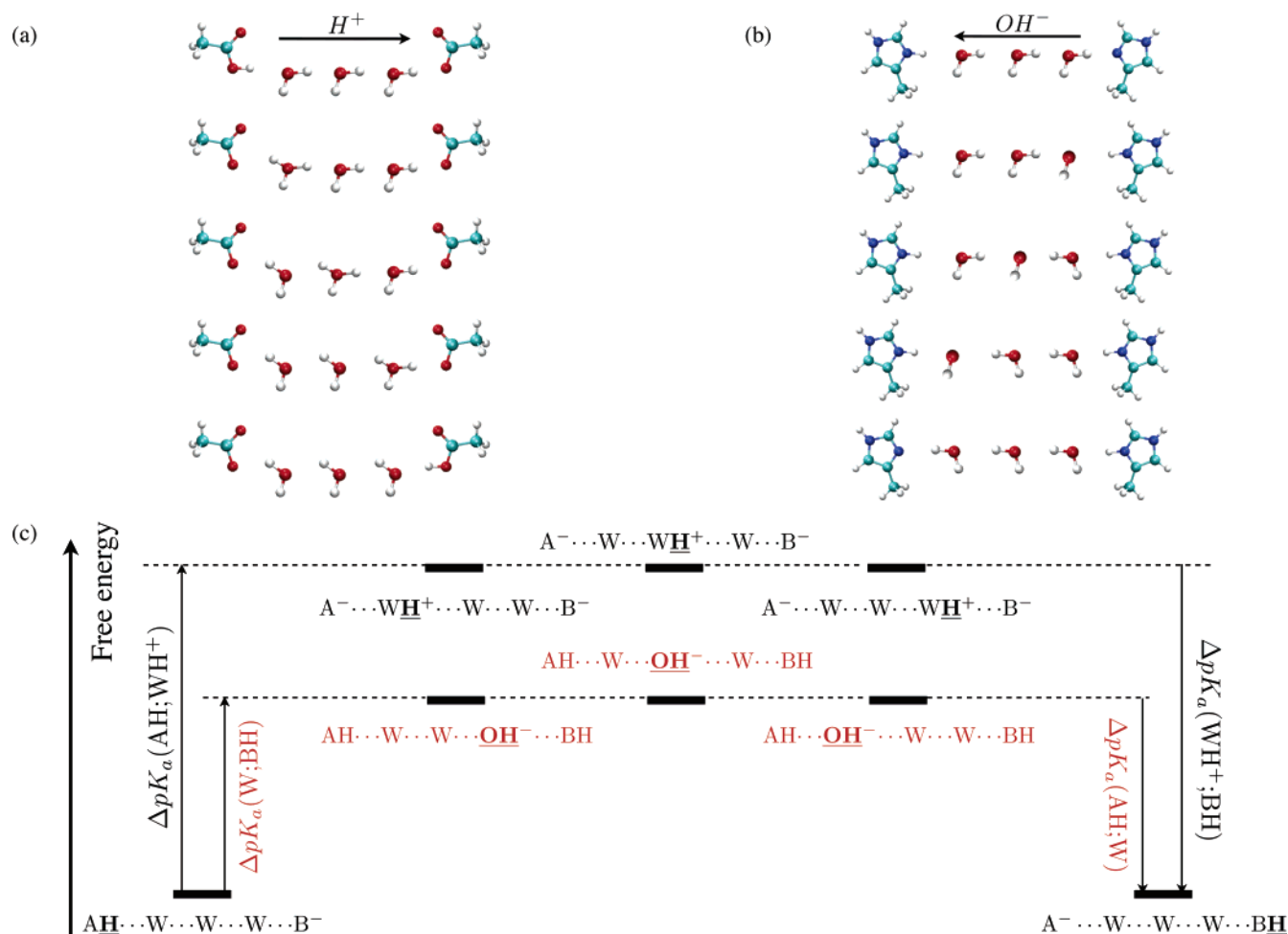


Figure 1. Molecular mechanism for the proton transfer between two molecules in solution depends critically on their pK_a values and electrostatics in the environment. (a) Classical Grotthuss mechanism that involves sequential proton hops from the donor to the acceptor; (b) hydroxide mediated mechanism that involves first protonating the acceptor and then a series of "proton hole" transfers toward the donor; (c) relative energetics of the Grotthuss and "proton hole" transfer mechanisms can be approximately estimated based on the pK_a values of the donor/acceptor groups (AH, BH) and water (W) as well as hydronium (WH^+). With the right pK_a combinations, (e.g., when both AH and BH have pK_a larger than 7), the "proton hole" transfer pathway can be energetically more favorable than the classical Grotthuss mechanism.

as hydroxide for water, have been rarely considered in either experimental or theoretical studies of proton-transfer reactions.¹³

Considering that protons are exchanged and charges are displaced, the intrinsic pK_a of the donor/acceptor groups and electrostatics due to the environment are expected to play a dominant role in determining the molecular details of the proton-transfer process; e.g., all accessible protonation states of the mediating groups, such as hydroxide for water, should be considered. Take a simple system in which a proton is transferred between two groups of similar pK_a values through a number of water molecules. As illustrated schematically in Figure 1c, when the end groups have acidic pK_a values ($\ll 7$), the classical Grotthuss mechanism is energetically favored. When the end groups have basic pK_a values ($\gg 7$), however, it is energetically more favorable to first deprotonate the last water molecule along the chain to protonate the final acceptor group, thus generating a solvated hydroxide; a set of subsequent "proton hole" transfers toward the initial donor group then takes place to complete the process (Figure 1b). Note that we prefer to call these "proton hole" transfers because this terminology describes any deprotonated molecule within the transfer pathway. When

the pK_a s of the end groups are close to 7, both the classical Grotthuss and the "proton hole" transfer mechanisms can contribute and which pathway dominates should depend sensitively on the electrostatic environment. In fact, these considerations suggest that for a reliable mechanistic description, the theoretical model has to treat the relative pK_a values of the donor, acceptor, and mediating groups in a balanced manner.

In the following, we use hybrid quantum mechanical/molecular mechanical (QM/MM) simulations^{14–17} (see Supporting Materials for details regarding the calibration of methods) to explicitly discuss the diverse molecular mechanisms for long-range proton transfers in realistic solution and enzyme systems. These simulations became practical only recently due to a number of useful technical developments made in our groups, which include a reliable treatment of electrostatics in the QM/MM framework,¹⁸ flexible samplings for long-range

(13) Eigen, M.; Hammes, G. G. *Adv. Enzymol.* **1963**, 25, 1.

(14) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, 11, 700.
 (15) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solution*; Wiley: New York, 1991.
 (16) Gao, J. In *Reviews in Computational Chemistry VII*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1995.
 (17) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, N.; Prat-Resina, X.; König, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M.; Cui, Q. *J. Phys. Chem. B* **2006**, 110, 6458.
 (18) Schaefer, P.; Riccardi, D.; Cui, Q. *J. Chem. Phys.* **2005**, 123, 014905.

proton transfer with highly nonlinear pathways¹⁹ and improvements of an approximate density functional theory.^{17,20–22} Although there are still limitations in the current methods, they are sufficient for establishing a computational framework to discuss the factors that dictate the detailed proton-transfer mechanism in complex molecular systems.

II. Systems and Simulation Methods

The main goal is to sample the relevant mechanisms of long-range proton-transfer reactions in complex molecular systems, such as solution and enzymes, and calculate the relevant energetics. The general simulation protocol employed is the hybrid quantum mechanical/molecular mechanical (QM/MM) approach^{14–17} that treats a small (typically ~100) number of atoms with QM while the rest with MM. The setup details are given below.

A. Solution Systems. To emphasize the connection between the pK_a values of donor/acceptor groups and the molecular mechanism of proton transfers, two proton exchange reactions in solution are studied. For each solution simulation, the setup is essentially the same; one solute and its conjugate base ($\text{CH}_3\text{COOH}/\text{CH}_3\text{COO}^-$; protonated and neutral 4-methyl-imidazole) are placed symmetrically about the origin and solvated in a 16 Å water sphere (~600 water molecules). The solutes and nearby water molecules are treated with SCC-DFTB^{20,21} (see Section II C for details) and the rest water with the TIP3P model.²³ For both systems, the donor and acceptor distance is approximately maintained throughout the simulations to be ~10 Å by applying positional constraints to a few solute atoms not directly involved in the proton transfer: all four carbons are constrained in the acetic acid simulations whereas the CE1, CD2, and the CB (methyl) carbons²⁴ are fixed for each 4-methyl-imidazole. To facilitate error analysis in the simulations (see Supporting Materials for details), several “half-reactions” that involve the proton exchange between a protonated/neutral 4-methyl-imidazole and water are also studied, in which a single solute is constrained to the center of the water sphere and a dummy atom is used to weakly constrain a nearby (5 Å from the donor/acceptor nitrogen) water molecule as a reacting partner.

Nonbonded interactions (electrostatic and van der Waals) for the inner region are calculated without cutoffs. The bulk electrostatics are considered via GSBP^{18,25} with an 18 Å inner region, which includes an additional 2 Å buffer region. The generalized reaction field matrix is evaluated using 400 spherical harmonics with an outer region dielectric constant of 80. Thermal collisions due to the bulk are included in the outer 3 Å of the 16 Å water sphere via Langevin dynamics²⁶ and the Langevin atom list is heuristically updated during the simulation. A 1 fs time step is used and the temperature is maintained at 300 K. Bonds involving hydrogen are constrained with the SHAKE algorithm²⁷ for MM waters and QM groups that are not explicitly involved in the proton transfer (see Section II C for details).

B. Carbonic Anhydrase II (CAII). Carbonic Anhydrase II (CAII) is a classical zinc enzyme that moderates respiration by catalyzing the interconversion between CO_2 and HCO_3^- . A large-body of kinetic studies have established that the rate-limiting step of the catalytic cycle involves a proton transfer from the zinc-bound water to a solvent accessible residue (His 64), under the high-buffer condition,²⁸ thereby

generating the zinc-bound hydroxide that acts as the nucleophile for CO_2 hydration. The zinc-bound water and His 64 are separated by at least ~8 Å,²⁹ thus water molecules in the active site have been proposed to facilitate the proton transfer via the classical Grothuss mechanism.²⁹ This water-wire based “proton hopping” mechanism has been largely accepted in the literature; the precise nature for the proton-transfer kinetic bottleneck and the dominant transfer pathway (if any), however, remain unclear.^{17,30–37} The relative small size of CAII and the rich experimental background make it an ideal system for the purpose of analyzing molecular details of proton transfer in realistic enzyme environment.

For the CAII simulations, the GSBP setup similar to that described in earlier publications^{17,18} is used. The crystal structure (2CBA³⁸) is centered with the zinc atom at the origin, and additional water molecules are added for proper solvation. The inner region has a radius of 20 Å. The zinc ion, its ligands, the final proton acceptor His 64, nearby water molecules, and Thr 200 are treated as QM (see Sect.II.C for details), while the rest of the system is described with the CHARMM 22 force field²⁴ (modified TIP3P for water³⁹). In the outer region, the dielectric interface between protein and bulk (dielectric constants of 1 and 80, respectively) is defined by the atomic radii of Roux and co-workers;⁴⁰ the inner region has a dielectric constant of one. The static field due to the outer region is evaluated with linear Poisson–Boltzmann using a focusing scheme that places a 56 Å cube of fine grid (0.4 Å) into a larger 132 Å cube of coarse grid (1.2 Å). The reaction field matrix is evaluated using 400 spherical harmonics.

During the MD simulations, all atoms in the outer region (along with some atoms at the edge of inner region) are held fixed and provide anchors for the system. As in the solution simulations, the inner region is further partitioned into Newtonian and Langevin regions:²⁶ all atoms between 16 and 20 Å are treated with Langevin MD while the rest with Newtonian MD; all protein atoms in the Langevin region are harmonically constrained with force constants corresponding to the B-factors in the PDB file. The Langevin atom list is updated heuristically, a 1 fs time step is used with SHAKE applied to groups not involved in the proton transfers, and the temperature is maintained at 300 K. Water molecules are kept within an 18 Å radius by a weak spherical boundary potential. Classical electrostatic interactions and van der Waals interactions are calculated without cutoffs, as in the solution simulations.

The side chain of His 64 was found to adopt two rotameric states in the high-resolution X-ray structure of CAII.³⁸ The Nδ atom in the “in” and “out” configurations is ~8 and ~11 Å from the zinc-bound water, respectively. It is generally assumed that the proton transfer occurs with only the “in” configuration,⁴¹ after which the protonated His 64 flips to the “out” configuration to release the proton to bulk solution, thus it is sufficient to focus on the results for the His64-in situation for the purpose of this work. An NOE restraint (with a force constant of 50 kcal/mol·Å⁻²) maintains the “in” conformation by acting on CG atom of His 64 only when the distance from the zinc atom is greater than 8.8 Å. In addition to the wild type simulations, simulations with

- (19) König, P.; Ghosh, N.; Hoffman, M.; Elstner, M. I. Tajkhorshid, E.; Frauenheim, T.; Cui, Q. *J. Phys. Chem. A* **2006**, *110* (Truhlar Issue), 548.
- (20) Elstner, M.; Porezag, D.; Jungnickel, G.; Elstner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.
- (21) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569.
- (22) Elstner, M. *Theo. Chem. Acc.* **2006**, *116*, 316.
- (23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (24) MacKerell, A. D., Jr. et al.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
- (25) Im, W.; Berneche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924.
- (26) Brooks, C. L., III; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159.
- (27) Ryckaert, J. P.; Cicciotti, G.; Berendsen, H. J. J. *Comput. Phys.* **1977**, *23*, 327.

- (28) Silverman, D. N. *Methods Enzymol.* **1995**, *249*, 479.
- (29) Eriksson, A. E.; Jones, T. A.; Liljas, A. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 274.
- (30) Aqvist, J.; Warshel, A. *J. Mol. Biol.* **1992**, *224*, 7.
- (31) Silverman, D. N. *Biochim. Biophys. Acta* **2000**, *1458*, 88.
- (32) Lu, D.; Voth, G. A. *Proteins: Struct., Funct., and Genet.* **1998**, *33*, 119.
- (33) Toba, S.; Colombo, G.; Merz, K. M. J. *J. Am. Chem. Soc.* **1999**, *121*, 2290.
- (34) Schutz, C. N.; Warshel, A. *J. Phys. Chem. B* **2004**, *108*, 2066.
- (35) Fisher, Z.; Hernandez Prada, J. A.; Tu, C.; Duda, D.; Yoshioka, C.; Govindasamy, A. H., L.; Silverman, D. N.; McKenna, R. *Biochemistry* **2005**, *44*, 1097.
- (36) Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 1071.
- (37) Smedarchina, Z.; Siebrand, W.; Fernandez-Ramos, A.; Cui, Q. *J. Am. Chem. Soc.* **2003**, *125*, 243.
- (38) Hakansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. *J. Mol. Biol.* **1992**, *227*, 1192.
- (39) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902.
- (40) Nina, M.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 5239.
- (41) Duda, D.; Tu, C.; Qian, M.; Laipis, P. J.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. *Biochem.* **2001**, *40*, 1741.

the His 64 mutated to an Ala are carried out (see Section III B for the motivation). The mutant has been shown to have a 20-fold reduction in catalytic efficiency in the absence of buffer,⁴² which corresponds to an increase in the barrier of ~ 2 kcal/mol. The setup is exactly the same as the wild-type enzyme with only the imidazole of His 64 substituted by a hydrogen; this alanine residue is no longer included in the QM region.

C. QM/MM Partitioning and Switching. For the solution simulations, the solute molecules and a cylinder of water molecules, defined about the line connecting the donor and acceptor atoms, are included in the QM region; the radius of the cylinder is taken to be ~ 5 Å. The QM/MM interaction consists of only electrostatic and van der Waals terms. Because water molecules are expected to diffuse rapidly, the QM and MM identities of the water molecules are allowed to switch during the simulation; i.e., at any given MD step, the distances of the furthest QM and closest MM molecules are compared and the identities are switched, if necessary; the frequency of switching is ~ 2 ps⁻¹. To avoid large fluctuations in the energy, the number of QM waters is defined at the start of the simulation and held fixed throughout; 15–30 water molecules (depending on the donor–acceptor separation) are included in the QM region. Two to four water molecules are sufficient to bridge the solute, donor, and acceptor atoms directly; the larger number of water molecules included in the QM region add flexibility to “where” and “how” the proton-transfer pathway is sampled.

For the CAII simulations, the QM region includes the zinc with its three histidine ligands (His94, His96, and His119) and one water, His 64, Thr 200, and all (~ 14) water molecules between the zinc bound water and His 64. For all histidine and threonine residues, a link atom⁴³ is placed between the α - and β -carbon atoms and the MM bonding terms are retained. As in the solution simulations, the QM and MM identities are allowed to switch to properly treat the exchange of water molecules between the active site region and the bulk; the frequency of switching events for the enzyme simulations is ~ 0.3 ps⁻¹.

For the QM level, the SCC–DFTB approach^{20,21} is chosen based on its overall balance of computational efficiency and accuracy.⁴⁴ Specifically, the standard second-order SCC–DFTB approach with improved hydrogen bonding interactions^{17,22} is used. As shown in previous studies, the hydrogen bonding correction is necessary to avoid unbalanced QM/MM and QM/QM interactions especially at the QM/MM boundary.¹⁷ As discussed in the Supporting Information, this parametrization of SCC–DFTB yields a balanced treatment for the proton affinities of the donor and acceptor groups in CAII, which is important for describing the corresponding reaction free energy. The errors in the proton affinity (PA) can be substantially reduced by using an improved version of SCC–DFTB including third-order terms,^{17,22} but this approach introduces errors in the PA difference between the zinc-bound water and His 64 in the CAII simulations. In the end, as long as the errors are understood, an accurate mechanistic picture can be extrapolated, as will be discussed below.

D. Potential of Mean Force (PMF) Calculations and Analysis.

To simulate proton transfers in the condensed phase, PMF calculations are carried out; this approach naturally includes fluctuations of water and protein atoms at the room temperature. To allow a flexible description of the proton-transfer process without an *a priori* biased transfer mechanism, a general coordinate introduced in our recent work¹⁹ is used:

$$\vec{\xi} = \sum_{i=1}^{N_H} \mathbf{r}^{H_i} - \sum_{j=1}^{N_X} w^{X_j} \mathbf{r}^{X_j} - \sum_{i=1}^{N_H} \sum_{j=1}^{N_X} f_{sw}(d_{X_i H_j})(\mathbf{r}^{H_i} - \mathbf{r}^{X_j}) \quad (1)$$

where \mathbf{r}^{atom} is the position of hydrogens (H) or heavy atoms (X), w^{X_j} is

the number of hydrogens associated with the heavy atom in its least protonated state as determined from the reactant and product; for example, in the autoionization of water, the donating water molecule has $w^{\text{O-donor}} = 1$ and the accepting water molecule has $w^{\text{O-accept}} = 2$. The switching function, $f_{sw}(d)$, allows the determination of bonds between heavy atoms X and hydrogen atoms H:¹⁹

$$f_{sw}(d) = 1/(1 + \exp[(d - r_{sw})/d_{sw}]) \quad (2)$$

where d is the distance between the heavy and hydrogen atoms; r_{sw} and d_{sw} are the parameters that enable this function to determine whether the proton is coordinated to the heavy atom or not. This coordinate locates the position of excess charge within the set of atoms involved in the proton transfer, which in turn can be utilized to define an one-dimensional reaction coordinate,

$$\zeta_R = \frac{|\mathbf{r}^D - \vec{\xi}|}{|\mathbf{r}^D - \vec{\xi}| + |\mathbf{r}^A - \vec{\xi}|} \quad (3)$$

where $\mathbf{r}^{D/A}$ is the position of the initial donor and ultimate acceptor atoms. When ζ_R is 0.0, 0.5, or 1.0, for example, the center of excess charge is at the donor, in the middle, or at the acceptor, respectively. As shown in the previous work, the ζ_R coordinate gives a rather flexible and reliable description for long-range proton transfers; e.g., adiabatic mapping calculations with ζ_R produced reaction paths very similar from minimum energy path calculations for proton transfers in the active site of CAII.¹⁹ The unique value of using a collective coordinate such as ζ_R is that the fluctuation between different types of water wires that connect the donor and acceptor groups^{17,18,32,33} is naturally taken account.

The PMF corresponding to this reaction coordinate is calculated using the umbrella sampling technique with 19 windows. Each window is sampled for ~ 120 ps for both solution and CAII. To ensure convergence for the CAII PMFs, three independent sets of simulations starting from different initial conditions (e.g., different length of the initial water wire connecting the donor and acceptor atoms) are used, which leads to a total simulation time on the order of ~ 7 ns for CAII. Data from different windows are collected and merged using the weighted histogram analysis method (WHAM)⁴⁵ approach. The convergence of PMFs is monitored by systematically removing blocks of trajectories from the WHAM calculations, which showed that all PMFs have converged within 1–1.5 kcal/mol. In addition, following the recent suggestion from Thiel and co-workers,^{46,47} the time series of the reaction coordinates was tested for trend in mean and variance (Kendall–Mann test for rank correlation) to ensure equilibration. For the data after equilibration, the data was coarse grained and tested for correlation (von Neumann test for serial correlation) and normality (Shapiro–Wilk test for normality). All of these tests were carried out at the 95% confidence level. The number of coarse grained intervals was 24 or larger in each case. These tests ensure equilibration and that the sampled data is characteristic of the ensemble.

To characterize the protonation states of heavy atoms involved in the proton transfer, the excess coordination as a function of the reaction coordinate, ζ_R , and their position relative to the reaction coordinate are analyzed. The excess coordination, $\nu_0^{X_i}$, of each heavy atom, X_i , is determined using an expression introduced by Chakrabarti et al.:⁴⁸

$$\nu^{X_i} = \left[\sum_{j=1}^{N_H} f_{sw}(d_{X_i H_j}) \right] - w^{X_i} \quad (4)$$

where $f_{sw}(d)$ and w^{X_i} is the same switching function and weight as described above; the coordination is normalized as follows,

- (42) Tu, C.; Silverman, D. N.; Forsman, C.; Jonsson, B.; Lindskog, S. *Biochem.* **1989**, *28*, 7913.
 (43) König, P. H.; Hoffmann, M.; Frauenheim, T.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 9082.
 (44) Elstner, M.; Frauenheim, T.; Suhai, S. *THEOCHEM* **2003**, *632*, 29.

- (45) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.
 (46) Kästner, J.; Thiel, W. *J. Chem. Phys.* **2006**, *124*, 234106.
 (47) Schiferl, S. K.; Wallace, D. C. *J. Chem. Phys.* **1985**, *83*, 5203.
 (48) Chakrabarti, N.; Tajkhorshid, E.; Roux, B.; Pomes, R. *Structure* **2004**, *12*, 65.

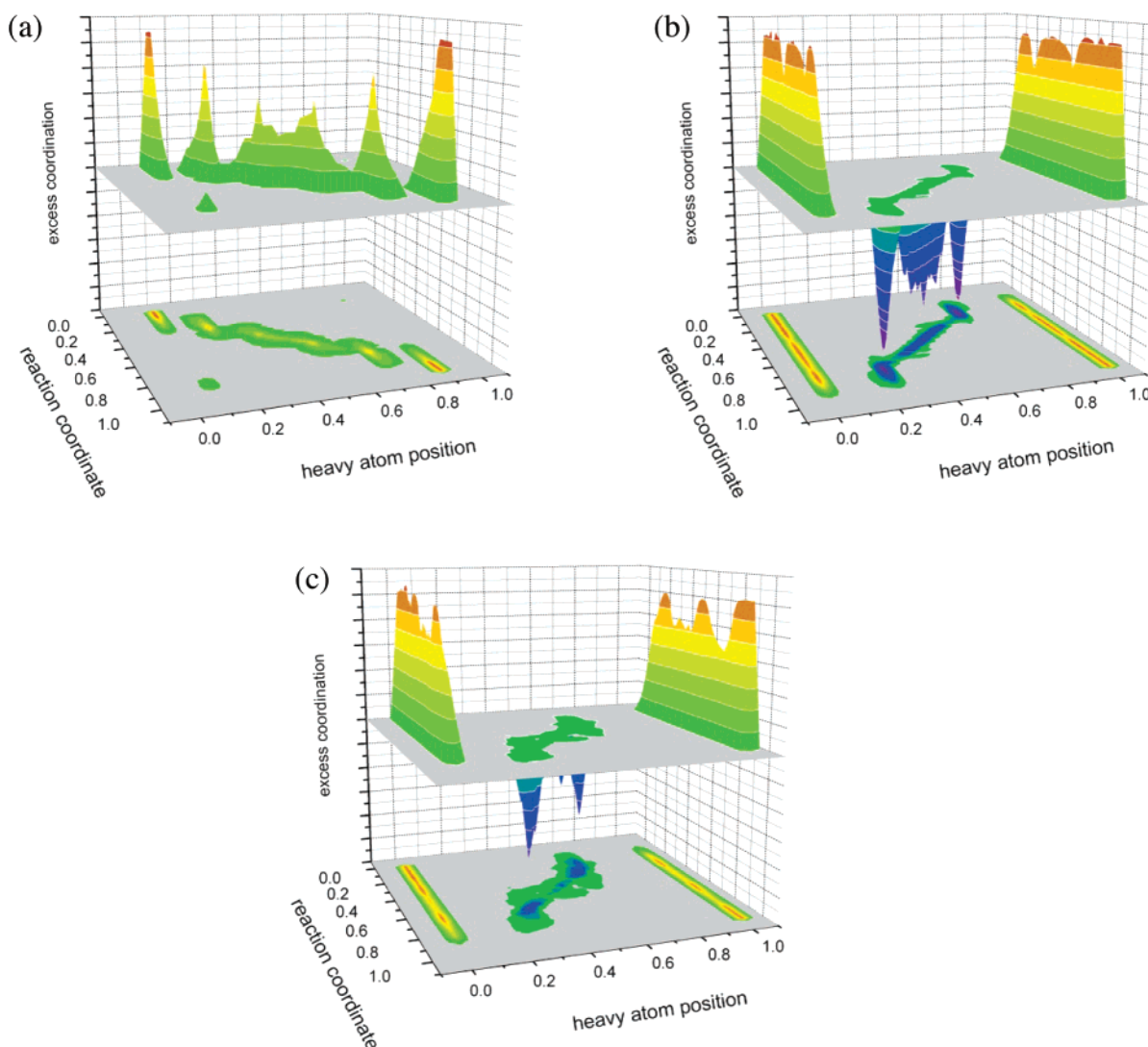


Figure 2. “Excess coordination” for the donor atom, acceptor atom and oxygen atoms in the mediating water molecules during the proton-transfer reactions. Each positive/negative peak in the plot corresponds to a protonated/deprotonated heavy atom in the course of the reaction. The plots are for proton transfers between (a) two acetic acids (one protonated, one deprotonated) in solution; (b) two 4-methyl-imidazole molecules (one protonated, one charge neutral) in solution with the modified (NHmod) nitrogen-hydrogen repulsive potential in SCC–DFTB; simulations with the standard nitrogen-hydrogen repulsive potential produce results very similar to (a) and are not shown (see text for discussions); (c) zinc-bound water and His 64 in the carbonic anhydrase II when His 64 adopts the “in” configuration.

$$\nu_0^{X_i} = \frac{\nu^{X_i}}{\sum_{j=1}^{N_X} \nu^{X_j}} \quad (5)$$

This normalization enables the qualitative descriptor (“excess coordination”, see Figure 2) to determine how the protonation states vary in space, relative to the reaction coordinate. The parameters d_{sw} and r_{sw} are adjusted for optimum differentiation between hydroxide and hydronium geometries ($r_{\text{sw}} = 1.25 \text{ \AA}$ and $d_{\text{sw}} = 0.04 \text{ \AA}$). For the position of the heavy atoms, we use a similar expression as ζ_{R} to measure the distance relative to the donor and acceptor atoms,

$$\zeta_{X_i} = \frac{|\mathbf{r}^{\text{D}} - \mathbf{r}^{X_i}|}{|\mathbf{r}^{\text{D}} - \mathbf{r}^{X_i}| + |\mathbf{r}^{\text{A}} - \mathbf{r}^{X_i}|} \quad (6)$$

E. Electrostatic Potential Calculations for Relevant Protonation States. Linearized Poisson–Boltzmann (LPBE) calculations are carried out using the PBEQ module in CHARMM to qualitatively evaluate the electrostatic potential for three protonation states of the CAII active site. The three protonation states are as follows: doubly deprotonated with a zinc bound hydroxide and deprotonated His 64 corresponding

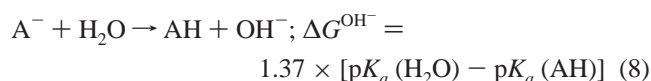
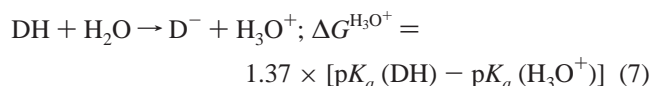
to the active state of the enzyme in the direction of CO_2 hydration; a singly protonated state with a zinc bound water and neutral His 64 corresponding to the active state of CAII in the dehydration direction; doubly protonated with a zinc bound water and protonated His 64. The doubly protonated and deprotonated states provide a description of the electrostatic potential seen by the hydroxide and hydronium proton-transfer pathways. The state that contains deprotonated zinc-water and protonated His 64 is not studied for its electrostatic potential despite its mechanistic relevance; this is because the purpose here is to study how the active site accommodates ions (hydronium or hydroxide) rather than neutral water.

To generate the three corresponding charge distributions, the solvated crystal structure (2CBA)³⁸ is used. The SCC–DFTB region containing the zinc with its ligands along with the protonated H64 is represented by SCC–DFTB Mulliken charges for the three protonation states with proper solvation of the active site. All water molecules not bound to zinc are then removed in the LPBE calculations. The protein dielectric boundary is defined with a water probe radius of 1.4 \AA and atomic radii of Roux and co-workers;⁴⁰ the dielectric constants of the protein and water are set to 4 and 80, respectively. The electrostatic potential

at 300 K with zero ionic strength is then computed with LPBE on a cubic grid with sides of 68.4 and 0.4 Å grid spacings.

III. Results and Discussions

In the current work, we describe the net exchange/transfer of a proton between two chemical groups (donor and acceptor) through a set of water molecules; the mechanistic issue of interest is the relative contribution of the classical Grotthuss mechanism and a hydroxide (or "proton hole") mediated pathway. As illustrated by Figure 1c, the energy penalties associated with the two pathways are expected to be strongly correlated to the pK_a of the relevant groups:



where the multiplication by 1.37 converts pK_a units to kcal/mol at room temperature. Note that by "energy penalty" we estimate the free energy cost for the rate-limiting event in a specific mechanism based on pK_a values for the relevant species; therefore, the barrier for the entire proton-transfer reaction following a given mechanism is expected to be correlated to such energy penalty. Generally speaking, the accuracy of calculated pK_a at a given level of theory for molecule *AH* depends on the calculated gas-phase proton affinity for *AH* and differential solvation free energy of *AH* and *A*[−]. Therefore, the environment may perturb the relevant pK_a values in eqs 7 and 8 such that the relative importance of the two pathways may, in principle, vary from one system to the other even if the same donor and acceptor groups are involved.

From eqs 7 and 8, it is clear that reliable simulations require balanced descriptions for the effective pK_a values of the donor, acceptor, water, and hydronium in the system of interest. Computing accurate pK_a values with all-atom QM/MM simulations⁴⁹ is challenging, especially for water and hydronium; therefore, our strategy focuses on understanding the systematic error in our methods (see Supporting Materials) and analyzing the physical interactions in the process of interest rather than spending huge efforts in parametrizing system-specific parameters. As illustrated below, although the relative contribution of Grotthuss and "proton hole" transfer pathways remains difficult to determine precisely, the current QM/MM simulations are sufficient for supporting the framework for analyzing factors that determine the relative contribution. We first investigate the solvent-mediated proton transfer between simple solutes in water and then discuss the interesting case of carbonic anhydrase.

A. Solution Cases. Acetic Acid. For a pair of conjugate acetic acids (one protonated and the other deprotonated) separated by 10 Å in solution, the electrostatic interaction between them is sufficiently screened and therefore their pK_a values are expected to be close to that of the isolated species, which is 4.8.⁵⁰ Therefore, it is expected that the classical Grotthuss mechanism is energetically more favorable than a hydroxide mediated "proton hole" transfer mechanism (Figure 1c). The plot of the

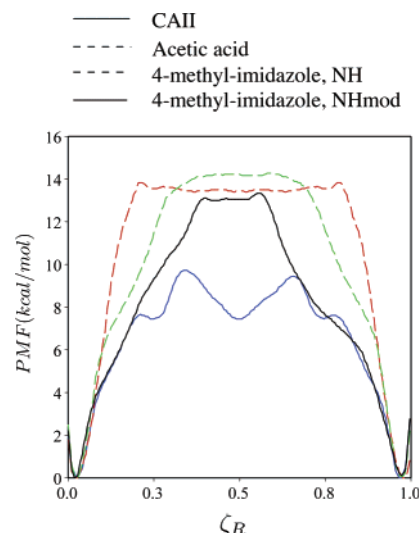


Figure 3. Calculated potential of mean force for the various proton-transfer reactions in solution and CAII based on SCC-DFTB/MM simulations; the "NH" and "NHmod" refer to the two sets of repulsive potential between nitrogen and hydrogen atoms in SCC-DFTB. The reaction coordinate is a collective coordinate that describes the progress of the proton transfer. For details, see Systems and Simulation Methods.

"excess coordination" for critical oxygen atoms during the simulation clearly (Figure 2a) shows positive peaks aligned diagonally; this confirms a classical Grotthuss mechanism with *sequential* proton hops. Considering that the corresponding rate-limiting process involves the proton transfer from the protonated acetic acid to the neighboring water (eq 7), the pK_a difference between the acetic acid and hydronium can be used to estimate the barrier to be ~9 kcal/mol, which is not far from the calculated value of 14.0 kcal/mol for the potential of mean force (PMF) barrier (Figure 3); the calculated value is overestimated due to the systematic error in the employed QM/MM method (see the Supporting Information). In addition, electrostatic contributions between the deprotonated solutes and the hydronium, in the barrier region, are expected to further lower the barrier by ~1 kcal/mol, this will be discussed next in the context of 4-methyl-imidazole.

4-Methyl-imidazole. For a pair of conjugate 4-methyl-imidazole molecules ("HisH⁺" and "His") separated by 10 Å in solution, the proton-transfer pathway is more interesting. Because the pK_a of the protonated 4-methyl-imidazole is ~7,⁵⁰ the classical Grotthuss and the "proton hole" mechanisms are energetically similar (Figure 1c). However, we note that the "proton hole" transfer pathway involves a larger degree of charge separation (hydroxide sandwiched between two protonated 4-methyl-imidazole) compared to the Grotthuss mechanism (localized hydronium between two neutral 4-methyl-imidazole) and therefore is expected to be energetically favored. In other words, the presence of the positively charged donor group (HisH⁺) stabilizes the generation of a nearby ion-pair (OH[−]↔HisH⁺), and it is this stabilization that makes the "proton hole" pathway more favorable than the Grotthuss mechanism; a simple continuum electrostatics model (see Supporting Materials) provides an estimate of 1.2 kcal/mol for the electrostatic stabilization, which is significant even if the slightly higher (by ~0.5 kcal/mol) mobility barrier of hydroxide relative to hydronium is taken into account.⁵¹

(49) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715.

(50) Lide, D. R. ed., *CRC Handbook Chemistry and Physics*, 85th ed.; CRC Press: Boca Raton, FL, 2005.

(51) Agmon, N. *Chem. Phys. Lett.* **2000**, *319*, 247.

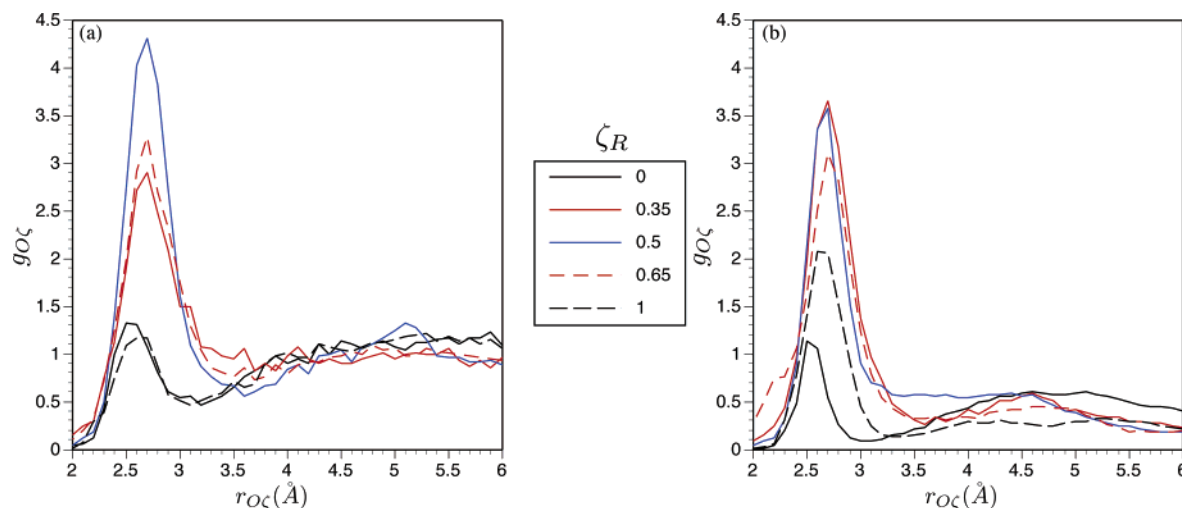


Figure 4. Radial distribution functions for the solvent oxygen around the “proton hole” for the proton-transfer reaction, between (a) two 4-methyl-imidazoles in solution and (b) the zinc-bound water and His 64 in CA II, at different stages in the process. The position of the “proton hole” is given by an expression similar to that for ξ in eq 1, with the opposite sign and weights of the donor and acceptor defined as the number of protons in the most protonated state relative the reactant and product (i.e., 2 for the zinc-bound water and 1 for Nδ in His 64).

The small value of the estimated electrostatic stabilization suggests that the sampled pathway in QM/MM simulations should be sensitive to the simulation details. As shown strikingly in Figure 2, the Grothuss and “proton hole” mechanism, respectively, is sampled as the dominant pathway with different SCC–DFTB parameters. With the standard NH repulsive potential,²⁰ the plot of the “excess coordination” for critical oxygen atoms during the simulation has the same feature as that in the case of acetic acids (Figure 2a), which points to the Grothuss pathway. With the NHmod repulsive potential,¹⁷ which was introduced to correct for proton affinity errors for sp^3 hybridized nitrogens, the “excess coordination” plot (Figure 2b) shows *negative* peaks aligned along the anti-diagonal, which is consistent with “proton hole” transfers toward the donor group.

Based on the error analysis discussed in details in the Supporting Information, the fact that different mechanisms are sampled with different sets of NH repulsive potential in SCC–DFTB is understood and expected from their effects on the pK_a differences. The calculated barriers of ~ 13 and 9.5 kcal/mol (Figure 3) for the Grothuss and “proton hole” pathways (based on different NH repulsive potential parameters), respectively, are within the expected error range from the value of 12.0 kcal/mol based purely on relative pK_a considerations. Although, the calculated barrier difference between the two pathways is smaller than the estimated systematic error for the QM/MM protocol used here and thus by itself does not prove that the “proton hole” pathway dominates, the fact that both barriers are close to the values expected based on eqs 7 and 8 supports the current framework of analysis. Therefore, the simulation results combined with the electrostatic argument presented above argue that the “proton hole” pathway is likely dominant for the proton exchange between 4-methyl-imidazole (and molecules with similar pK_a values) in solution. This prediction can be tested with full QM or QM/MM simulations with higher level QM methods as well as ultrafast spectroscopic measurements similar to that in ref.¹⁰

B. Carbonic Anhydrase II. Wild-Type (WT) CAII. The estimated classical barrier from the PMF calculations for the WT CAII is ~ 13 kcal/mol (Figure 3), which is ~ 3 kcal/mol

higher than the experimental estimate of 9.5 kcal/mol (with a maximal rate of $10^6 s^{-1}$).²⁸ Considering that including zero-point energy and proton tunneling typically reduces the barrier height by $2\text{--}3$ kcal/mol for proton transfers at 300 K, the agreement between experiment and calculation is encouraging. Moreover, it is satisfying that the experimentally observed thermoneutrality of the reaction is also reproduced by the current simulations; the calculated and measured values are $+0.1$ and -0.4 kcal/mol (the experimentally extracted donor/acceptor pK_a value is 6.8 ± 0.3 and 7.1 ± 0.3 , respectively⁴¹), respectively.

Strikingly, the “excess coordination” plot shown in Figure 2c clearly demonstrates that the sampled proton-transfer mechanism in CA II is largely dominated by the “proton hole” transfers, as indicated by the negative peaks along the anti-diagonal. Examination of snapshots from the simulations (see Figure 5a for an example) shows that the hydroxide is stabilized by many polar interactions in the active site, which involve not only water molecules but also sidechains of polar amino acids such as Thr 200, Asn 62, Asn 67, and Gln 92. Therefore, although hydroxide or “proton hole” transfers have not been discussed in any previous studies of CAII, it is clear that the active site of this enzyme is capable of accommodating the negatively charged ion (also see below).

This somewhat unexpected result might, however, be an artifact of the simulation protocol used. Indeed, the above discussion of the 4-methyl-imidazole case highlights the trend that the “proton hole” transfer mechanism is artificially favored over a hydronium pathway when the NHmod set of parameters is used in SCC–DFTB, due largely to the unbalanced description of proton affinity of water, imidazole and solvation of hydroxide (see the Supporting Information). Therefore, it is important to estimate the energetics for a hydronium pathway in CAII and compare the barrier difference to the magnitude of the error inherent in the current SCC–DFTB/MM simulations if any conclusive statement is to be drawn. Unlike the case of proton exchange between 4-methyl-imidazole and its conjugate acid in solution, switching to a different NH repulsive potential is not a good option because the reaction in CAII is no longer symmetric, and therefore, the calculated free energy will also be largely perturbed. Instead, we chose to calculate the energetics

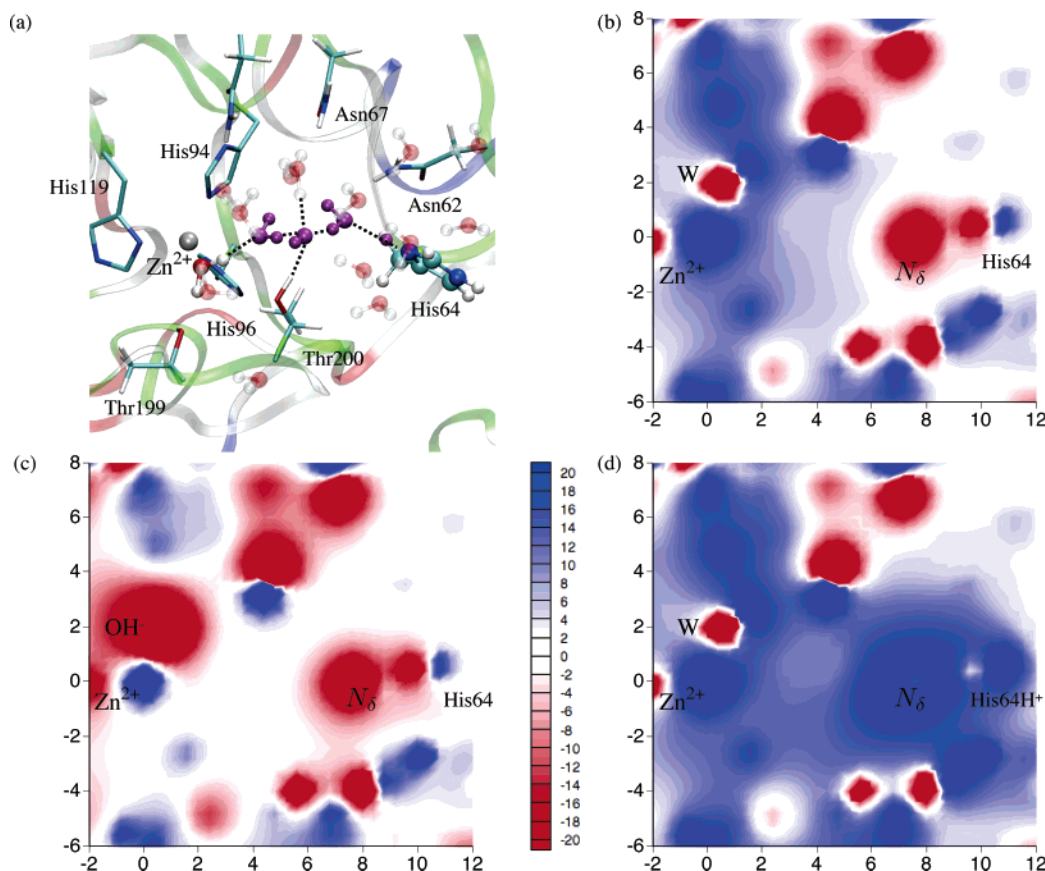


Figure 5. Snapshot and electrostatic potentials of the carbonic anhydrase II active site. (a) A snapshot from the SCC–DFTB/MM umbrella sampling simulation that illustrates the hydrogen bonding interactions that stabilize the hydroxide; the chain of water molecules that mediate the “proton hole” transfer are colored purple. (b–d) Calculated electrostatic potentials in the plane formed by the zinc ion, oxygen in the zinc-bound water and N_δ in His 64; the zinc ion is placed in the origin. The three plots are generated with different combinations of the protonation states of the zinc-bound water and His 64. (b) $Zn^{2+} \cdot H_2O/His64$; (c) $Zn^{2+} \cdot OH^-/His64$; (d) $Zn^{2+} \cdot H_2O/His64H^+$. Note that no other explicit water molecules are included in the Poisson-Boltzmann calculations for the electrostatic potentials.

ics for the proton transfer from the zinc-bound water to active site waters, for the H64A mutant, with the final acceptor water located ~ 5 Å away. Without His64 or any external “rescue” agent as the acceptor, the proton transfer must follow the classical Grotthuss mechanism. Since the site of mutation is at least 8 Å away from the zinc-water site, the energetics for this proton transfer is expected to be very similar to the Grotthuss mechanism in WT CAII. Experimentally, the overall rate of proton transfer in H64A was observed to drop by a factor of 20, in the absence of buffer, relative to the WT;⁴² a possible source of this rate reduction is discussed below.

The H64A Mutant and Grotthuss vs “Proton Hole” Pathways in CAII. The calculated PMF indicates that the initial proton transfer in H64A is up hill by ~ 21 kcal/mol. Although this value appears much higher than the calculated barrier of the “proton hole” transfer mechanism found in the WT simulations, there are substantial systematic errors associated with the current SCC–DFTB/MM simulations. The thermoneutrality of the WT PMF (Figure 3) reveals the similarity in the pK_a s of the zinc-bound water and His 64 with the NHmod repulsive potential in SCC–DFTB. Considering the experimental fact that the pK_a of His 64 is not shifted substantially between solution and the CAII active site, the energetics of proton transfer between a protonated 4-methyl-imidazole and water in solution should be comparable to the proton transfer discussed above for H64A, if the solvation errors associated with solution and enzyme simulations are similar. For the proton exchange

between protonated 4-methyl-imidazole and water, the formation of hydronium is overestimated by ~ 9 kcal/mol; if one assumes that the systematic error transfers from solution to the enzyme systems, which is commonly done in the simulation field,¹⁵ then the “corrected” estimate for the proton-transfer barrier in H64A is $\sim 21 - 9 = 12$ kcal/mol. On the other hand, the proton transfer from water to a neutral 4-methyl imidazole is underestimated by ~ 4 kcal/mol; therefore, with a similar correction the barrier found in the WT PMF would increase from ~ 13 to 17 kcal/mol. While this supports the Grotthuss mechanism for CAII, the true character of the transfer remains murky as transferability of the error from solution to the enzyme environment is unclear. In fact, comparing the solvent structure around the “proton hole” during the reaction in solution (Figure 4a) and CAII (Figure 4b) clearly shows that the “proton hole” in the barrier region ($\zeta_R \approx 0.5$) is substantially less solvated in CAII relative to solution. Because a large portion of the systematic error in the current simulations of pathways involving hydroxide is the oversolvation of the hydroxide ion (see Supporting Materials), it is logical to expect that the systematic error is reduced in the enzyme relative to solution although a quantitative measure is difficult.

Does the existence of considerable systematic errors in our simulations mean the current study offers little new mechanistic insight? We strongly argue otherwise because the current work, for the first time, brings up the possibility that “proton hole” transfer may make an important contribution in addition to the

classical Grotthuss mechanism. Considering the relatively small *absolute* error associated with the calculated energetics for the “proton hole” pathway (see Supporting Information), the current simulations demonstrate that this alternative to the Grotthuss mechanism is not strongly disfavored in CAII and in fact has energetics reasonably consistent with experimental measurements. The difficulty in pinning down the relative energetics for the hydroxide and hydronium mediated pathways only highlights the challenge associated with simulating reactive events involving “delocalized” species in the condensed phase, such as hydronium and hydroxide. This is a major challenge not only for semiempirical level of QM/MM methods but also higher-level QM/MM simulations, especially considering the importance of extensive sampling. Therefore, instead of putting huge efforts into “fixing” SCC–DFTB/MM for the problem by developing system-specific parameters, we make several observations to supplement the QM/MM results and argue that the “proton hole” transfer mechanism likely makes an important, if not dominant, contribution in the WT CAII. Verification of this prediction requires further theoretical as well as experimental studies; in this regard, we note that probing mutants with shifted pK_a values for the zinc-water and the proton acceptor group is of particular relevance, although it is also possible that the dominant mechanism in the mutant is different from the wild type system due to the shifts in the pK_a values.

First, the pK_a values for the initial donor ($Zn^{2+}\cdot H_2O$) and final acceptor (His 64) are experimentally determined to be ~ 7 ; thus, both the “proton hole” and Grotthuss mechanisms are likely to contribute if the enzyme active site resembles bulk water. Which mechanism dominates depends largely on the electrostatics in the active site, which dictates whether a positively (hydronium) or negatively (hydroxide) charged species is better stabilized. Considering that the biological function of CA II is to accommodate the negatively charged bicarbonate, it is plausible that the active site is configured through evolution to favor a pathway involving a negatively charged species rather than a positively charged species. In Figure 5, the electrostatic potential of the active site is shown in three different combinations of protonation states of the zinc-bound water and His 64. As long as the zinc-bound water is charge neutral (Figure 5b,d), the electrostatic potential in the region between the zinc site and His64 is largely positive, especially when His 64 is protonated (Figure 5d). Therefore, the electrostatics of the active site is fully compatible with the generation of a negatively charged hydroxide and subsequent “proton hole” transfers. When the zinc-bound water is deprotonated (Figure 5c), the electrostatics between the zinc site and the neutral His64 are nearly neutral thus do not particularly favor the generation of hydroxide or disfavor the presence of hydronium. The sensitivity of potential to the protonation state is consistent with the *reversibility* of catalysis for CAII; i.e., when the zinc-bound water is protonated, the positive potential primes the enzyme for the dehydration of bicarbonate.

Second, it was established experimentally that the rate of proton transfer is reduced by a factor of 20, in the absence of buffer, in the H64A mutant.⁴² The EVB analysis of Warshel et al.³⁰ argued that the rate for the classical hydronium pathway in the WT CAII is dictated by the first proton transfer from the zinc-bound water to the next water molecule. Therefore, to rationalize the drop in the overall rate in the H64A mutant, it is

necessary to argue that the proton transfer to water molecules at the surface of CAII is substantially slower than those in the active site. Although this may be justified by a higher reorganization penalty associated with the former process, detailed analysis has not been available. If the “proton hole” mechanism is dominant in the wild type CAII as we suggest, a simple alternative explanation can be presented: the “proton hole” mechanism is more favorable in the wild type but not possible in the H64A mutant. In fact, the 20-fold difference in rate might provide an estimate for the difference between the hydronium and hydroxide mediated mechanisms in the wild type CAII. Which explanation is valid can be probed by detailed simulation of proton transfer from the active site to the surface of the protein, which is currently underway.

IV. Concluding Remarks

Although the quantitative results presented here for specific systems are subject to limitations associated with the simulation techniques and much study remains to thoroughly analyze the large body of kinetic data available for CAII and its mutants, the calculations here established a framework emphasizing that long-range proton transfers in solution and enzyme can exhibit very different molecular details depending on the pK_a values of the participating groups and electrostatic interactions with the environment. Therefore, we argue that the classical Grotthuss mechanism that has dictated the analysis of long-range proton-transfer processes is likely oversimplified and should be supplemented with the possibility of “proton hole” transfers being equally important. Although the role of hydroxide transfer has been eluded to in previous discussions of proton transport in bacteriorhodopsin,⁵² the current analysis provides the first direct argument with energetic considerations concerning a “proton hole” pathway. The results also further highlight the diverse roles of water in biomolecules and extend the list of “functional waters” to include not only protonated water¹¹ but also hydrated hydroxide.

In addition to biomolecules, proton transfers in nanomaterials are also expected to exhibit rich behaviors considering the interplay between polar interactions and hydrophobic confinements. We encourage experimental identification of hydroxide signatures in these systems using state-of-the-art spectroscopic techniques. The difficulty with a more quantitative computational analysis highlights the challenge associated with simulating reactive events involving “delocalized” species such as hydronium and hydroxide in the condensed phase, and calls for the development of theoretical methods that can incorporate all accessible protonation states of key species in a computationally efficient manner.

Acknowledgment. The research discussed here have been partially supported from the National Science Foundation (MCB-0314327, CAREER-0348649) and the National Institutes of Health (R01-GM071428-01). Q.C. also acknowledges a Research Fellowship from the Alfred P. Sloan Foundation. Computational resources from the National Center for Supercomputing Applications at the University of Illinois are greatly appreciated.

(52) Facciotti, M. T.; Rouhani, S.; Glaeser, R. M. *FEBS Lett.* **2004**, *564*, 301.

Supporting Information Available: Detailed discussions regarding relevant benchmark calculations as well as error analysis for the current SCC-DFTB/MM method are available.

The full list of authors for ref 24 is also given. This material is available free of charge via the Internet at <http://pubs.acs.org>.
JA065451J