HHS Public Access

Author manuscript

J Phys Chem B. Author manuscript; available in PMC 2020 August 26.

Published in final edited form as:

J Phys Chem B. 2020 July 09; 124(27): 5696–5708. doi:10.1021/acs.jpcb.0c03615.

Understanding and Tracking the Excess Proton in Ab Initio Simulations; Insights from IR Spectra

Chenghan Li[†], **Jessica M. J. Swanson**[‡]

[†]Department of Chemistry, Chicago Center for Theoretical Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States

[‡]Department of Chemistry, Biological Chemistry Program, and Center for Cell and Genome Science, The University of Utah, Salt Lake City, Utah 84112, United States

Abstract

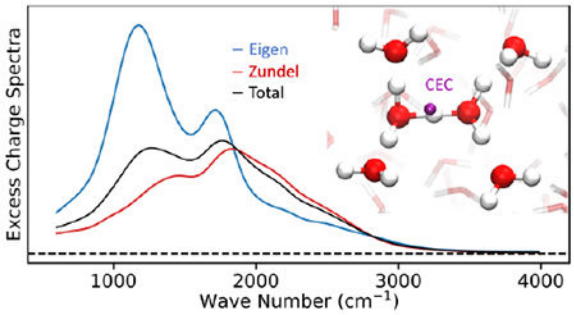
Proton transport in aqueous media is ubiquitously important in chemical and biological processes. Although ab initio molecular dynamics (AIMD) simulations have made great progress in characterizing proton transport, there has been a long-standing challenge in defining and tracking the excess proton, or more properly the center of excess charge (CEC) created when a hydrogen nucleus distorts the electron distributions of water molecules in a delocalized and highly dynamic nature. Yet defining (and biasing) such a CEC is essential when combining AIMD with enhanced sampling methods to calculate the relevant macroscopic properties via free energy landscapes, which is standard practice for most processes of interest. Several CEC formulas have been proposed and used, but none have yet been systematically tested or rigorously derived. In this paper, we show that the CEC can be used as a computational tool to disentangle IR features of solvated excess proton from its surrounding solvent, and in turn, how correlating the features in the excess charge spectrum with the behavior of CEC in simulations enables a systematic evaluation of various CEC definitions. We present a new definition of CEC and show how it overcomes the limitations of those currently available both from a spectroscopic point of view and from a practical perspective of performance in enhanced sampling simulations.

Graphical Abstract

Corresponding Author: Jessica M. J. Swanson (j.swanson@utah.edu).

Supporting Information

Additional figures containing the mCEC bias forces acting on atoms of Asp in water in metadynamics, the vibrational density of states (VDOS) of the ground state coefficients c_i^2, and the comparison between mCEC and rCEC on the time evolution of proton disassociation coordinate ξ of Asp in water in SMD. Details of the mathematical proof of the proton donor identity independency of rCEC and methods of computing the VDOS of c_i^2.



1. Introduction

Understanding proton transport (PT) is of central importance in many domains of biology and condensed-phase chemistry. The hydrated excess proton is often characterized by two limiting structures known as the Eigen $(H_9O_{4+})_1$ and Zundel cations $(H_5O_{2+})_2$, although it is more properly described as a dynamic, delocalized charge distribution that fluctuates around these limiting structures.[3, 4] Solvated protons diffuse significantly faster than other cations with similar charge and size in water due to their unique Grotthuss hopping mechanism, wherein the excess positive charge is transported through water molecules by rearranging covalent and hydrogen bonds.[4, 5]

When considering PT in complex systems such as weak acids or biomolecules, there are typically protonation and deprotonation rare events where intermediates are separated by barriers of 5 to 20 kcal/mol. In these cases, enhanced sampling is essential to capture the PT process in accessible simulation times. Popular enhanced sampling methods, including umbrella samplin,[6, 7] metadynamics,[8–12] and adaptive bias force,[13, 14] involve the use of bias forces on one or more degrees of freedom of the system often called collective variables (CVs). A fundamental requirement for a CV is its differentiability in order for the bias forces to be decomposed onto individual atoms. Due to the delocalized nature of the positive excess charge and long time scale of many PT events, a CV that tracks the position of the excess proton is needed. The concept of a center of excess charge (CEC) was developed in the 90's for this reason.[15,16] One might assume that integrating the net electron distribution (accounting for periodic boundaries) would yield the CEC. However, it is impossible to separate instantaneous charge fluctuations due to water-water interactions from the charge transfer induced by the excess proton. As pointed out by Konig et al.,[17] some early definitions did not accurately describe the location of the excess proton due to the contamination of irrelevant water molecules. Compared to these early definitions, Konig's mCEC (modified CEC) has shown robustness in several test cases and was applied to model PT in several systems.[18–23] More recently, Pezeshki and Lin proposed another CEC definition called the proton indicator that was employed as collective variable in umbrella sampling of proton permeation through a hydrophobic carbon nanotube.[24] Although the mCEC and the proton indicator have been demonstrated in simulating explicit PT, they both involve *ad hoc* switching functions and parameters that lack clear physical meaning. Thus, a solid foundation and a rigorous definition of the CEC are still missing.

IR spectroscopy is a powerful technique to study the solvation structures and transport mechanism of the excess proton in acid solutions[25–33]. However, IR signals of condensed phase acid systems come from all of the proton-water complexes involved in the delocalized charge defect, and thus the experiments themselves are usually not sufficient to determine the molecular origin of the vibrational modes that are being excited and probed. Although the "bulk water part" of acid solution is typically subtracted off to get the difference spectrum, as shown in Figure 1A, the coupling and delocalization of vibrational modes as well as degeneracy of multiple motions in a certain frequency range makes the interpretation of IR spectra controversial.[34, 35] Recent 2D-IR experiments by Elsaesser and co-workers suggested that a double minimum model best describes the proton transfer mode potential energy surface (PES),[36] while similar experiments by Tokamokoff and coworkers claimed a PES with a single minimum.[37] They also have divided opinions on the lifetime of the proton complex: Tokamokoff claimed a lifetime over 480-fs,[30] while Elsaesser observed a sub-100-fs population decay of the proton transfer mode first excited state.[36]

The CEC, as described above and detailed below, describes the motions of the solvated proton but is by design insensitive to bulk water motions. This inspired us to investigate whether or not the CEC can, in addition to serving as a collective variable in enhanced sampling, also be used as a computational tool to study the IR spectroscopic properties of the protonated water complex in the aqueous environment under finite temperature with no need of extracting clusters[34, 35, 38–42], solute-solvent partitioning[43], independent simulations of pure water[39, 44, 45] nor the harmonic approximation for normal mode analysis[35, 39]. And in turn, can the spectrum decode the resulting collective motions, such that the spectra of different CEC definitions can inform us about the dynamical behaviors and performance of that CEC in enhanced sampling simulations?

However, defining the CEC is not trivial, as the positive charge due to the presence of an excess proton is distributed across many water molecules, each with an altered electron distribution, and over multiple solvation shells. Moreover, this distribution is highly dynamic with the partially covalent hydrogen bonding network shifting on the femtosecond timescale. A successful PT reaction is considered to occur when the dominant central hydronium-like water molecule changes identities, and is thought to happen on the picosecond timescale. According to the Eigen-Zundel-Eigen (EZE) proton transport mechanism validated in simulations, the most probable proton acceptor around the hydronium (the "special pair partner") oscillates between the three waters surrounding the central hydronium on the 25-100 femtosecond timescale during a relatively long-lived Eigen-like period called the "special pair dancc"[3]. Intermittent with this, there are Zundel-like periods in which one "special pair partner" is favored and the excess positive charge bounces back in forth between two waters, albeit in a highly asymmetric manner. A successful PT then occurs when the special pair partner in Zundel-like dynamics takes over and becomes the new central hydronium-like molecule, which is thought to depend on coordinated hydrogen bonding in the second solvation shell.[46, 47]

Throughout the process, asymmetry is favored and the distribution of excess charge at any given time predominantly involves 6 water molecules and most resembles an asymmetric Eigen-like/Zundel-like structure in which one special pair partner involves more charge

transfer than the other two and including the two water molecules that special pair partner is hydrogen-bonded to in the second solvation shell.[46] Thus, the symmetric Eigen and Zundel cations are somewhat limiting in our ability to describe or conceptualize the process as a whole. What is perhaps a better descriptor of the state of the excess proton is captured in its dynamics. In simulations, there are clear periods of Eigen-like dynamics featuring the special pair dance, and periods of Zundel-like dynamics in which one special pair partner is singled out. To properly track this highly dynamic behavior requires a rigorous definition of the CEC that resolves the identity of proton acceptor on-the-fly and enables smooth transitions among the most probable acceptors.

These complexities also impose challenges on modeling PT processes in computer simulations. One solution is the *ah initio* molecular dynamics (AIMD)[48] where electronic structure is treated explicitly and thus the changing bonding topology is captured. AIMD can, in principle, achieve a high order of accuracy in computing the potential energy surface (PES) and thus has been widely used to study aqueous PT and solvations[5, 48–55]. However, the most accurate AIMD simulation of acid solutions may require electronic structure methods to determine intemuclear forces that correctly describe the dispersion, charge transfer and many-body effects[56–58] in hydrogen bond networks, as well as possibly an appropriate treatment of nuclear quantum effect (NQE).[59–62] These methods are generally too expensive for condensed phase simulations. Thus, the commonly used methods are more affordable generalized gradient approximation (GGA) level of density functional theory (DFT), which is known to produce over-structured condensed phase water and to under-estimate the diffusion of water.[63] Even at these levels, AIMD simulations still have a high computational cost that limits the time and space scales attainable. The time scale problem, for all simulations, can be addressed by combination with advanced enhanced sampling methods, which again underlines the importance of a well-behaved CEC definition.

An alternative solution is multi-scale reactive molecular dynamics (MS-RMD), that evolved from the multi-state empirical valence bond (MS-EVB) method and has been long developed by the Voth group.[64–73] The MS-RMD methodology has been benchmarked and demonstrated to describe proton solvation and transport in various aqueous and biomolecular systems.[20, 21, 71, 73–78] In the MS-RMD framework, the system's ground state PES is defined as a linear combination of 'valence bond' states in which electrons are considered to be localized in valence bonding orbitals and thus representative of specific bonding topologies. A Hamiltonian matrix describes state potential energies on diagonal terms and coupling between states on off-diagonal terms, thereby capturing the delocalized nature of the charge defect. Redefinition and weighting of the state contributions at every time step captures the dynamics of the charge defect and changing bonding topologies. The CEC position in MS-RMD is naturally defined as a linear combination of the charges in each valence bond state.[79] However, the MS-RMD Hamiltonian is only directly available in MS-RMD simulations and thus the MS-RMD CEC cannot be used in AIMD or other reactive models.

In this work, we introduce a new definition of the CEC, namely rCEC, which is designed to reproduce the MS-RMD CEC charge transfer behavior while still being available in AIMD simulations. We demonstrate that the spectrum of the CEC produces almost all of the

signature features of the acid solution revealing that the collective motions of the protonated water complex and the surrounding hydrogen bond networks are all encoded in the motion of the center of excess charge. The two most commonly used existing AIMD CECs (mCEC and the proton indicator) are analyzed in detail and compared with the new rCEC, both from a spectroscopic point of view and in simulations. We show limitations of mCEC and the proton indicator reflected by their spectra and further validated in simulations, and how these limitations are overcome by the new rCEC.

2. Methods

2.1 Theory of MS-RMD and MS-RMD CEC

In the MS-RMD framework, the ground state of the molecular system $|\psi\rangle$ is expanded on a basis set where each basis state $|i\rangle$ corresponds to a distinct bonding topology (Figure 2), i.e.

$$|\psi\rangle = \sum_{i=1}^{N} c_i |i\rangle \tag{1}$$

where the c_i's are the expansion coefficients and N is the total number of basis diabatic states. The c_i's are obtained on-the-fly over the course of simulation by solving the eigenvalue problem using the Hamiltonian $\mathbf{H} = \sum_{ij} h_{ij}|i\rangle\langle j|$, such that

$$\mathbf{H}\boldsymbol{c} = E_0\boldsymbol{c} \tag{2}$$

where E_0 is the ground state energy and $\boldsymbol{c} = \{c_i\}$ is the ground state coefficient vector. The diagonal terms h_{ii} are expressed in two contributions, a molecular mechanics forcefield term h_{ii}^{MM} and a diabatic correction, h_{ii}^{CORR} for shifting the arbitrary baseline energy common to classical force fields. The off-diagonal term h_{ij}, which provides the mechanism for the system to transition between basis states, is typically a geometric function with tunable parameters. The details are well-illustrated elsewhere.[71, 73]

In each RMD diabatic state, the system is governed by classical mechanics where the bonding topology is well-defined and hydrogen bonds are purely described by non-bonded interactions. In such description, the excess proton is simply identified according to the bonding topology and thus the CEC is simply the center of charge (COC) of the moiety that owns the excess proton in entirety. The resulting definition of center of excess charge is then a linear combination of the center of charge (COC) in all MS-RMD states,

$$\boldsymbol{r}_{\mathrm{CEC}} = \sum_{i=1}^{N} c_i^2 \boldsymbol{r}_i^{\mathrm{coc}}. \tag{3}$$

Here, c_i is the ground state coefficient and the COC, $\mathbf{r}_i^{\mathrm{coc}}$, is essentially the "center of excess charge" in a given bonding topology of the i-th RMD state:

$$\boldsymbol{r}_i^{\mathrm{coc}} = \frac{\sum_{J \in Q_i} |q_J| \boldsymbol{r}_J}{\sum_{J \in Q_i} |q_J|} \tag{4}$$

where subscript J goes over all atoms in the excess proton species (e.g., all the four atoms in a hydronium or the carboxylic hydroxyl for a protonated amino acid).

2.2 Definition of rCEC

Noting that the non-trivial parts in eq 3 are the c_i's, one can easily compute CEC in AIMD as long as those coefficients are known. Thus, defining CEC in AIMD is essentially calculating c_i's for each given atomic configuration. As c_i gives the magnitude of how much each COC contributes to the CEC, c_i's is strongly correlated to some appropriate measure of the charge transfer (e.g., from the hydronium to its solvation waters). An energy decomposition analysis (EDA) study of the charge delocalization of an excess proton in water indicated that the following asymmetry coordinate (Figure 2) is a good measure of the charge-transfer energy:

$$\delta_{IJK} = r_{JK} - r_{IK} \tag{5}$$

, where I, J are indices of the two proton-donor/acceptor oxygens while K is the index of hydrogen between them.

We found an exponential function of this delta value best fits how much charge is transferred from the central hydronium to its first solvation shell water (Figure 2):

$$\frac{c_j^2}{c_i^2} = f_{\mathrm{ct}}(\delta_{IJK}) \tag{6}$$

Here, I indicates the HB donor atom in MS-RMD state $|i\rangle$, K is the proton to be transferred and J is the HB acceptor atom in MS-RMD state $|j\rangle$;

$$f_{\mathrm{ct}}(x) = \begin{cases} e^{-kx} & x \le r_1 \\ a_0 + a_1 x + a_2 x^2 + a_3 x^3 & r_1 < x \le r_2 \\ 0 & x > r_2 \end{cases} \tag{7}$$

, where the damping term brings the exponential to exact zero. Interestingly, we note that eq 6 indicates the charge transfer factor c_j^2/c_i^2 is the Pauling bond order (BO) ratio of two waters and our fitted parameter agrees to some extent with what Pauling suggested for BO_{80} (Table 1).

We further assume that this rule holds for charge transfer from the first solvation shell molecules to the second solvation shell and so on. With this assumption, the excess charge on the central hydronium follows the rule (eq 6) to propagate layer by layer along hydrogen bonding networks. The hydrogen bond donor/acceptor pairs are resolved on-the-fly by a state searching algorithm, which finds every possible bonding topology based on geometric criteria. In this work, the EVB3 state searching algorithm with three solvation shells was

used as its good energy conserving behavior implies an exhaustive search is achieved.[73] First, every hydrogen is assigned to the closest oxygen. The oxygen with three assigned hydrogens (in the case of a single excess proton) is designated the central pivot oxygen (the most hydronium-like water). Then hydrogen bonds trace all potential bonding topologies within given number of solvation shells. The coefficient of the central hydronium is determined by exploiting the normalization condition, $|\boldsymbol{c}|^2 = 1$:

$$c_1^2 = \frac{1}{1 + \sum_{i=2}^{N} c_i^2 / c_1^2} \tag{8}$$

In the case of an amino acid solvated in water, the proton transfer reaction between the acid and a water is not symmetric as it is for water and water. Thus, it requires an extra parameter δ_0 to take this asymmetry into account. We use

$$f_{\mathrm{ct}}(x) = \begin{cases} e^{-k(x-\delta_0)} & x \le r_1 \\ a_0 + a_1 x + a_2 x^2 + a_3 x^3 & r_1 < x \le r_2 \\ 0 & x > r_2 \end{cases} \tag{9}$$

to compute how much excess charge is transferred from the protonated amino acid to its first solvation shell while keeping the charge transfer behavior between water molecules the same. The coefficients in the polynomials in both f_{ct}'s are chosen to ensure f_{ct}'s C_1 smoothness:

$$\begin{aligned} f_{\mathrm{ct}}(r_1 -) &= f_{\mathrm{ct}}(r_1 +) \\ f'_{\mathrm{ct}}(r_1 -) &= f'_{\mathrm{ct}}(r_1 +) \\ f_{\mathrm{ct}}(r_2 -) &= 0 \\ f'_{\mathrm{ct}}(r_2 -) &= 0 \end{aligned} \tag{10}$$

The four conditions determine the polynomial coefficients uniquely:

$$\begin{aligned} A &\triangleq \exp(-k(r_1 - \delta_0))/(r2 - r1)^3 \\ a_0 &= A \cdot r_2^2\left(3r_1 + kr_1^2 - r_2 - kr_1 r_2\right) \\ a_1 &= A \cdot r_2\left(6r_1 + 2kr_1^2 - kr_1 r_2 - kr_2^2\right) \\ a_2 &= -A \cdot \left(3r_1 + kr_1^2 + 3r_2 + kr_1 r_2 \right. \\ &\quad \left. -2kr_2^2\right) \\ a_3 &= A \cdot (2 + kr_1 - kr_2) \end{aligned} \tag{11}$$

2.3 Relation between CEC and IR spectrum

IR spectroscopy is a powerful tool of studying molecular events involving vibrational motions. It reflects how the system's total dipole responds to an applied external electric field and thus it can be computed from the dipole-dipole time correlation function according to the fluctuation-dissipation theorem. In classical MD, the system dipole is expressed as the total dipole of nuclei point charges as the electronic DOFs are not explicitly treated:[81]

$$\boldsymbol{\mu} = \sum_I q_I \boldsymbol{r}_I \tag{12}$$

For a proton in water, the total atomic charge can be decomposed into two contributions: the "intrinsic" charge, which is the partial charge on each atom in the absence of excess protons, and the excess charge, which is the additional atomic charge that arises from the delocalization and polarization of the protonic charge defect. Thus, the total dipole moment is expressed as:

$$\boldsymbol{\mu} = \sum_I \left(q_I^{\mathrm{ex}} + q_I^{\mathrm{in}}\right)\boldsymbol{r}_I = \boldsymbol{\mu}^{\mathrm{ex}} + \boldsymbol{\mu}^{\mathrm{in}}. \tag{13}$$

The IR absorption is thus decomposed into four terms:

$$\begin{aligned} I(\omega) &\propto \int \langle \dot{\boldsymbol{\mu}}(0)\dot{\boldsymbol{\mu}}(t)\rangle e^{-i\omega t}\mathrm{d}t \\ &= \int \left\langle \dot{\boldsymbol{\mu}}^{\mathrm{ex}}(0)\dot{\boldsymbol{\mu}}^{\mathrm{ex}}(t)\right\rangle e^{-i\omega t}\mathrm{d}t \\ &+ \int \left\langle \dot{\boldsymbol{\mu}}^{\mathrm{ex}}(0)\dot{\boldsymbol{\mu}}^{\mathrm{in}}(t)\right\rangle e^{-i\omega t}\mathrm{d}t \\ &+ \int \left\langle \dot{\boldsymbol{\mu}}^{\mathrm{in}}(0)\dot{\boldsymbol{\mu}}^{\mathrm{ex}}(t)\right\rangle e^{-i\omega t}\mathrm{d}t \\ &+ \int \left\langle \dot{\boldsymbol{\mu}}^{\mathrm{in}}(0)\dot{\boldsymbol{\mu}}(t)\right\rangle e^{-i\omega t}\mathrm{d}t \end{aligned} \tag{14}$$

Interestingly, the general form of the CEC for each of the CEC definitions studied herein is a summation of the product of atomic excess charge and position:

$$\boldsymbol{r}_{\mathrm{CEC}} = \sum_I q_I^{\mathrm{ex}} \boldsymbol{r}_I \tag{15}$$

, where $\sum_I q_I = 1$ holds for the reason that total excess charge induced by a proton is unitary. Thus, the CEC is actually the dipole of excess charges up to a factor of unit charge $\bar{e}$, i.e.

$$\boldsymbol{\mu}^{\mathrm{ex}} = \bar{\mathrm{e}}\boldsymbol{r}_{\mathrm{CEC}} \tag{16}$$

This leads to the fact that the excess charge contribution to IR spectra is proportional to the Fourier transformation of CEC velocity-velocity correlation function

$$I^{\mathrm{ex}}(\omega) \propto \int \langle \boldsymbol{v}_{\mathrm{CEC}}(0)\boldsymbol{v}_{\mathrm{CEC}}(t)\rangle e^{-i\omega t}\mathrm{d}t \tag{17}$$

The formula coincides with the "vibrational density of states" (VDOS) of the center of excess charge. However, we should point out that the center of excess charge is not a real particle but instead a collective motion of hydrogen bonding networks, including many vibrations and charge fluctuations in the protonated water cluster and thus the VDOS of the CEC has a more profound physical meaning than the VDOS of real atoms. In contrast to the partial charges of atoms, the excess charges are expected to decay along increasing solvation shells of the excess proton and thus separate the protonated water complex IR absorption from the total absorption of the system.

2.4 rCEC spectrum decomposition

2.4.1 Solvation shell decomposition—Decomposing rCEC into solvation shell contributions is straightforward from eq 15:

$$\begin{aligned}\boldsymbol{r}_{\mathrm{CEC}} &= \boldsymbol{r}_D + \sum_{I \in \{\mathrm{H_3O}\}} q_I^{\mathrm{ex}} \boldsymbol{r}'_I + \sum_{I \in \left\{1^{\mathrm{st}}\mathrm{Shell}\right\}} q_I^{\mathrm{ex}} \boldsymbol{r}'_I \\ &+ \sum_{I \in \left\{2^{\mathrm{nd}}\mathrm{Shell}\right\}} q_I^{\mathrm{ex}} \boldsymbol{r}'_I + \sum_{I \in \left\{3^{\mathrm{rd}}\mathrm{Shell}\right\}} q_I^{\mathrm{ex}} \boldsymbol{r}'_I \triangleq \boldsymbol{r}_D + \boldsymbol{r}_{\mathrm{CEC,H_3O}} + \boldsymbol{r}_{\mathrm{CEC},\,1^{\mathrm{st}}\mathrm{Shell}} \\ &+ \boldsymbol{r}_{\mathrm{CEC},\,2^{\mathrm{st}}\mathrm{Shell}} + \boldsymbol{r}_{\mathrm{CEC},\,3^{\mathrm{rd}}\mathrm{Shell}}\end{aligned} \tag{18}$$

where $\boldsymbol{r}_D$ is the position of the donor atom (i.e. oxygen atom of the most probable hydronium) and $\boldsymbol{r}'_I = \boldsymbol{r}_I - \boldsymbol{r}_D$ is the relative positions of atoms with respective to the donor atom. The first solvation shell was defined as the direct hydrogen bond acceptors of the H_3O motif. The second shell was defined as the direct hydrogen bond acceptors of the first shell and so on. Similar to eq 17, the solvation shell dependent spectrum was computed by the Fourier transform of CEC velocity auto-correlation function of each solvation shell component.

2.4.2 Eigen/Zundel decomposition—At each timestep, the most probable hydronium oxygen (O_0) and the special pair oxygen (O_{1x}) were identified based on the two largest c_i^2. The whole trajectory was then partitioned into non-reactive segments separated by timesteps where O_0 is changed. In each segment, the number of unique O_{1x} identities (n) was counted. If $n = 3$, special pair dance happens in this segment and we assigned this segment as Eigen. If $n = 1$, only one special pair is favored and thus is assigned as Zundel. For segments with $n = 2$, it could be part of a concerted proton hopping or partial special pair dance, making any assignment controversial, so we keep this case independent instead of merging into Zundel or Eigen. The resulting population of each assignment is Zundel:Eigen:unassigned = ~1:5:1. Considering the $n = 2$ spectrum is closer to Zundel (Figure 1B), the Eigen/Zundel ratio is roughly 5:2, which corresponds to a roughly 0.55 kcal/mol free energy difference, agreeing well with Figure 2 in ref [82].

According to the Wiener-Khinchin theorem, the excess charge spectrum (eq 17) can be rewritten as

$$I^{\mathrm{ex}}(\omega) \propto \left| \int \boldsymbol{v}_{\mathrm{CEC}}(t) e^{-i\omega t} dt \right|^2 \tag{19}$$

This inspired us to use short time Fourier transform (STFT) to compute a time-resolved excess charge spectrum:

$$I^{\mathrm{ex}}(t, \omega) \propto \left| \int \boldsymbol{v}_{\mathrm{CEC}}(\tau) W(\tau - t) e^{-i\omega t} \mathrm{d}\tau \right|^2 \tag{20}$$

, where $w(t)$ is a Tukey window with a length of 175 fs. The choice of the length is a balance between time-resolution and Fourier transform accuracy. Other reasonable choices do not

change the result qualitatively. Along with the Eigen/Zundel assignments at each time, the spectra computed by eq 20 were independently averaged according to the value of n.

2.5 Parametrization of rCEC

Reference charge transfer factors c_j^2/c_i^2 are obtained from diagonalizing the Hamiltonian matrix in MS-RMD simulations. We intend to find the optimal parameters k and δ_0 in f_{ct} functions to reproduce the reference MS-RMD charge transfer behaviors. Instead of minimizing the CEC position discrepancy, the following object function was used

$$\chi^2(\{\theta\}) = \left\langle |\Delta \boldsymbol{c}|^2 \right\rangle = \left\langle |\boldsymbol{c}_{\mathrm{CEC}} - \boldsymbol{c}_{\mathrm{MS-RMD}}|^2 \right\rangle \tag{21}$$

, where $\{\theta\}$ are the parameters to be fit, which is k in f_{ct} for hydronium-water or k and δ_0 for weak acids; brackets indicate the average in MS-RMD ensemble; $\boldsymbol{c}_{\mathrm{CEC}}$ indicates the coefficients calculated by the algorithm described in Section 2.2 and $\boldsymbol{c}_{\mathrm{MS-RMD}}$ is the reference coefficients from the MS-RMD model.

For proton and water, the MS-RMD ensemble was sampled for 500-ps of one excess proton and 256 waters (simulation details in Section 2.7.2). The rCEC parameters for the amino acid were fitted in the ensemble sampled from the 9 windows ranging from 1.25 Å to 2.50 Å of MS-RMD umbrella sampling. We used a two-step strategy to find the optimized parameters. In the first step, only the first solvation shell was considered, for which a deterministic least square fitting can be used. The obtained parameter value obtained was then fed as the initial guess for a second step Powell's method that included the second and the third solvation shells and for which the first shell parameters depended[83]. The second step optimization was conducted with an in-house script which utilizes GNU parallel[84] for acceleration. The final resulting parameter in step 2 is rather similar compared to those in step 1 (see Table 1), which validates our assumption in Section 2.2 that charge transfer along multiple shells follows the same exponential rule (eq 6).

2.6 Other CEC definitions

The mCEC uses a fermi function $f_{sw}(x) = 1/[1 + \exp((x - r_{sw})/d_{sw})]$ to switch between 0 and 1 (nonbonding to bonding) for each heavy atom-hydrogen atom pair, effectively counting the number of hydrogens each heavy atom is bonded to. The definition of mCEC reads[17]

$$\boldsymbol{r}_{\mathrm{mCEC}} = \sum_{I \in \{H\}} \boldsymbol{r}_I - \sum_{J \in \{X\}} w_J \boldsymbol{r}_J - \sum_{I \in \{H\}} \sum_{J \in \{X\}} f_{sw}(r_{IJ})(\boldsymbol{r}_I - \boldsymbol{r}_J) \tag{22}$$

where $\{H\}$ is the collection of all hydrogen atoms and $\{X\}$ is all heavy atoms that accept or donate protons. The weight w_J is defined as the number of protons boned to atom J in its reference state so that a heavy atom has close-to-zero contribution to the position of CEC if it is in its deprotonated state.

The definition of the proton indicator is as follows (details can be found elsewhere[85])

$$\boldsymbol{r}_{\mathrm{Ind}} = \frac{1}{g_{\mathrm{Ind}}}\left(\boldsymbol{r}_D + \sum_{J \in S_X(r_{\mathrm{LIST}})} \sum_{I \in B_J} g(\rho_{IJ})\boldsymbol{r}_I\right) \tag{23a}$$

$$g_{\mathrm{Ind}} = 1 + \sum_{J \in S_X(r_{\mathrm{LIST}})} \sum_{I \in B_I} g(\rho_{IJ}) \tag{23b}$$

$$\rho_{IJ} = \frac{\boldsymbol{r}_{DI} \cdot \boldsymbol{r}_{DJ}}{|\boldsymbol{r}_{DJ}|^2} \tag{23c}$$

where $S_X(r_{\mathrm{LIST}})$ is the set of heavy atoms within r_{LIST} distance from the donor atom, D; B_J is the set of hydrogens bonded to heavy atom J; g is a function that switches from 0 when ρ_{MJ} is sufficiently small to 1 when $\rho_{MJ} > 0.5$.

2.7 Simulation Details

2.7.1 AIMD simulations—All AIMD simulations were conducted using the quickstep module in CP2K[86] with the hybrid Gaussian and plane wave (GPW) scheme[87] coupled with a modified version of PLUMED 2[88]. The protonated water system consists of 1 excess proton and 128 water molecules in a 15.66 Å × 15.66 Å × 15.66 Å box simulated in the NVE ensemble. The Gaussian and plane wave (GPW) method was employed and a plane wave cutoff of 400 Ry was used to expand the electron density. The Goedecker-Teter-Hutter pseudopotentials[89] were used to describe core electrons. The orbital transformation method was used to optimize the wave function at each step with a convergence criteria of 10_{-7} a.u.. The electronic structure was described by BLYP/TZV2P level DFT. In addition, an empirical D3 correction[90] and recently developed EDS correction[91] were employed. The choice of this level of AIMD is a reasonable balance between accuracy and computational cost and the EDS corrected version of BLYP is one of the best affordable ab initio methods for capturing water and hydrated proton solvation structure and dynamics.[82, 91, 92] We also emphasize that our CEC definition does not rely on the underlying PES and thus can be applied to higher level of DFT simulations. A timestep of 0.5 fs was used to integrate the system. The total simulation time was ~225 ps.

The aspartic acid system consists of 1 neutral Asp solvated in 141 waters in a 16.175 Å × 16.175 Å × 16.175 Å box simulated at 300 K in the NVT ensemble with a timestep of 0.5 fs. The BLYP/TVZ2P with D3 correction was used. In order to describe the proton disassociation process, we define a collective variable ξ as the distance between CEC and its closest carboxyl oxygen:

$$\xi = -\frac{1}{\beta}\log(e^{-\beta(d_1 - \bar{d})} + e^{-\beta(d_2 - \bar{d})}) + \bar{d} \tag{24}$$

, where $d_1 = |\boldsymbol{r}_{\mathrm{CEC}} - \boldsymbol{r}_{\mathrm{O_1}}|$ and $d_2 = |\boldsymbol{r}_{\mathrm{CEC}} - \boldsymbol{r}_{\mathrm{O_2}}|$ are the distances between the CEC and carboxyl oxygens and $\bar{d} = \frac{(d_1 + d_2)}{2}$. A sufficient large $\beta = 40$ was chosen to ensure the expression gives precise $\min(d_1, d_2)$.

The initial structures for steered MD (SMD) were generated by extracting configurations and velocities every 25 fs in a 4.5 ps-long trajectory at 300 K after a ~3 ps long equilibration. The bias potential used in SMD was

$$U(t) = \frac{1}{2}\kappa \cdot (\xi(\boldsymbol{r}^{3N}(t)) - \xi_0(t))^2. \quad (25)$$

The force constant κ was 80 kcal/mol/Å$_2$. The restraint center was

$$\xi_0(t) = \begin{cases} \xi(\boldsymbol{r}^{3N}(0)) + \frac{1.5\,\mathring{A}}{25\text{ fs}}t, & t < 25\text{ fs} \\ \xi(\boldsymbol{r}^{3N}(0)) + 1.5\,\mathring{A}, & t \geq 25\text{ fs} \end{cases} \quad (26)$$

A Gaussian hill height of 1 kcal/mol and addition pace of 25 fs used in metadynamics followed previous QM/MM metadynamics work93 while hill width of 0.2 Å was used and a bias factor $\gamma = 16$ was chosen based on an expected ~9 kcal/mol barrier for aspartate deprotonation[94].

2.7.2 MS-RMD simulations—Taking the advantage of its efficiency, MS-RMD simulations were used for umbrella samplings, and unbiased simulations to generate ensembles for fitting. The simulations were conducted in LAMMPS with an in-house MS-RMD package, RAPTOR,[95] coupled with PLUMED 2. The HCl solution system consists of 1 excess proton, 1 chloride and 256 water molecules in a 15.66 Å × 15.66 Å × 15.66 Å box. The MS-EVB 3.2 proton-water model was used.[72] Four NVE trajectories were initiated from independent configurations generated from a NVT sampling at 300 K using Nose-Hover chain of length 3. A timestep of 0.5 fs was used and the total simulation time was ~2.7 ns.

The glutamic acid system consists of 1 neutral glutamic acid and 1001 waters, resulting in a 31.075 Å × 31.075 Å × 31.075 Å box. A timestep of 1.0 fs was used. The temperature was controlled at 310 K with a fixed box volume. The MS-RMD parameters were taken from references [96] and [94]. The collective variable was defined as the distance between CEC and the center of mass of the carboxyl group. A 0.25 Å spacing of CV value was used for almost all the umbrella sampling windows while 0.1 Å was used for windows near the deprotonation transition state, resulting in 40 windows in total. All windows were run for over 1.5 ns.

3. Results and Discussion

3.1 rCEC spectrum decodes experimental IR spectrum for excess proton

By noting that CEC is actually the dipole moment of excess charges (eq 15), we compute the excess charge spectra (eq 17) of rCEC and the other two most commonly used CECs (mCEC and proton indicator) from AIMD to study the associated IR signals. All the CECs produce the acid continuum 1000 cm$_{-1}$-3200 cm$_{-1}$ (Figure 1C), which is the signature absorption of hydrated protons. The experimental IR spectrum[30, 35] features three broad peaks located at around 1200 cm$_{-1}$, 1750 cm$_{-1}$, and 2800 cm$_{-1}$. As shown in Figures 1A

and 1C, the rCEC spectrum not only outlines the overall shape of the experimental spectrum but also matches the 1200 and 1750 cm_{-1} frequency well, which are commonly assigned to the proton transfer mode (PTM) and the flanking water bending of the excess proton, respectively[44, 97]. This indicates that the CEC, originally designed for tracking the position of excess proton, is able to capture the motion of the excess proton as well as the associated stretching and bending motions of the protonated water complex.[35, 43] The rCEC spectrum decays to zero at around 3200 cm_{-1}, which is the same position as the difference spectrum zero point, right before the bulk-like water O–H stretching band (~3200-3700 cm_{-1}), validating the use of excess charge to compute IR spectrum of the protonated water complex without being contaminated by bulk water frequencies. The only deviation of the rCEC spectrum from experiment is that the intensity in the frequency range 2000 to 3200 cm_{-1} gradually decreases to zero as a function of frequency without showing a peak at 2800 cm_{-1}. Noting that the excess charge also dissipates along solvation shells, this decaying intensity suggests that this frequency range is due to the hydrating water that is spectroscopically different from bulk-like water, but more remote from the central excess proton. This can be verified by decomposing the CEC spectrum into contributions from different solvation shells (Figure 1D), which clearly shows the spectral absorption of each shell is gradually blue-shifted when going from the 1_{st} to the 3_{rd} solvation shell. The H_3O_+ core exhibits clear frequencies commonly assigned as the PTM and flanking water bending, verifying that the central H_3O_+, is intimately involved in those vibrations. The first solvation shell shows a broad peak at around 2100 cm_{-1}, possibly being the contributor to the continuum of 2000-2600 cm_{-1} in the difference spectrum. Interestingly, both the second and the third solvation shells absorb at around 2800 cm_{-1}, the other signature signal of the experimental spectrum. However, the averaged excess charge for 2_{nd} and 3_{rd} shell atoms is 8×10^{-4} and 2×10^{-5} in unit charge respectively, which are both too small for this IR signal to be visible in the total CEC spectrum. This is further verified by plotting the vibrational density of states for the populations of the dominant valence bond states (Figures S3A & B), which show that the 2800 cm_{-1} is clearly captured by the 4_{th} valence bond state.

State-of-art 2D-IR experiments[25, 30, 33, 36, 37, 98] have advanced our thinking on the solvation and transport of the excess proton in aqueous solution, but it has been a long-standing challenge to clearly correlate the spectroscopic observations to molecular vibrations. One of the debates has been whether or not the dominant hydrated proton structure more resembles the Zundel or Eigen cation. As suggested by Swanson and Simons,[46] the Zundel/Eigen species are perhaps better distinguished by their unique dynamical behaviors than by static geometric criteria. We thus decomposed the CEC spectrum according to Zundel/Eigen dynamics making use of the short-time Fourier transform (STFT) to get time-resolved CEC spectrum (eq 20). The AIMD trajectories were first partitioned into Zundel or Eigen segments according to the number of unique special pair identities (n) and the spectra for these segments are separately averaged. As shown in Figure 1B, both the Zundel ($n = 1$) and Eigen ($n = 3$) components exhibit very clear PTM and flanking water bending features. Noticeably, the Zundel has a stronger PTM that is slightly red-shifted while Eigen has a weaker and blue-shifted PTM. The same frequency shifting of Zundel/Eigen happens for the flanking water bending. The remaining unassigned spectrum component ($n = 2$) shows a mixture of Zundel and Eigen features with a PTM located between the Eigen/Zundel PTMs

while a flanking water bending closer to Zundel, implying it is a mixture of controversial Zundel and Eigen. These components merge into the total spectrum PTM and flanking water bending forming single peaks at frequencies roughly in the middle of the corresponding Zundel and Eigen frequencies. The Zundel and Eigen are more spectroscopically distinguishable at PTM with Zundel ~280 cm_{-1} lower than Eigen, but less separable for the flanking water bending with Zundel ~170 cm_{-1} lower. The proximity of the Zundel/Eigen frequencies suggests lasers in IR experiments centered at the middle frequencies of the PTM and flanking water bending modes[33, 37, 98] likely excite both the Zundel and Eigen species.

3.2 CEC spectra reveal their behaviors in simulations

Proton transport in an aqueous environment involves the stretching and bending of the protonated water complex as well as concerted motions of hydrogen bonding network.[3, 47, 99] Biasing the CEC in enhanced sampling is essentially enhancing the fluctuations in the direction of the collective motions related to PT. An optimal PT CV is expected to drive the system following the same path in phase space as PT naturally happens at equilibrium. As the collective motions of proton solvation and transport are captured in the IR spectrum, we expect the optimal CEC to match it (at least peak locations if not intensity). In this light, deviations from the difference spectrum can serve as a measure of the quality of a CEC definition. We thus investigated how the spectroscopic features of a CEC relate to that CEC's performance in simulations.

As shown in Figure 1C, although the first two peaks are also produced by mCEC and the proton indicator, there are some distinct features compared to rCEC. The spectrum of proton indicator shows non-zero absorptions in the full frequency range computed (0 cm_{-1}-30000 cm_{-1}). This spectral feature can be understood by noting the proton indicator's intrinsic discontinuity, as shown by its time series in Figures 3A and 3B. The x coordinate of the proton indicator, as an example of its xyz position, jumps over 1 Å multiple times within the 150-fs time segment, while smooth changes are observed for other CEC definitions (Figures 3A and 3C). In enhanced sampling, a continuous and differentiable collective variable is required to ensure a reliable resulting free energy profile; thus, the proton indicator should be revised to a differentiable form if it is going to be used in biased simulations.

In addition to a spectroscopic point of view, the proton indicator's discontinuity can be directly understood by examining its definition. The definition of proton indicator (eq 23) depends on the proton donor oxygen identity, which undergoes discrete changes introducing discontinuity in the CEC. This is well illustrated by Figures 3D and 3E where we see a large displacement in proton indicator's position within a 0.5 fs separation.

However, the proton indicator is continuous and differentiable for a fixed donor identity due to the smooth functions used in its definition, which can be checked by the smooth regions between jumps in Figure 3A or 3B. In other words, proton indicator hops between the smooth surfaces corresponding to different donor atom identities and the discrepancy in each surface results in its discontinuity (Figure 3B). It's worth noting that the rCEC also requires a donor atom to initiate the state searching algorithm and thus its differentiability relies on the agreement between the surfaces, which is a property of rCEC that can be understood mathematically (see SI). It is also validated by Figure 3C where rCEC trajectories computed

using different fixed pivot indices all overlay with the one computed with dynamically changing identity.

The most notable feature of mCEC is its non-vanishing absorption over 3200 cm_{-1} corresponding to the water stretching band (Figure 1C). This observation inspired us to examine if mCEC includes bulk water bond oscillations in addition to the motion of hydrated proton. In order to obtain meaningful statistics, we conducted an ensemble of steered MD simulations of deprotonating an aspartic acid solvated in water (Figure 4A). The distance between CEC and closest carboxylic oxygen (ξ) was employed as the CV (eq 24). The SMD protocol we used can be divided into two phases. The CEC was first pulled 1.5 A away from its initial position in 25 fs, which was implemented by applying a moving harmonic restraint on the collective variable ξ. Then the restraint was held at the final position of the first phase for another 75 fs. This mimics a standard procedure for using SMD to prepare initial configurations for umbrella sampling simulations. Figure 4B shows the magnitude of bias force acting on water atoms further than 6 Å from the carboxylic oxygens in each SMD run as a function of time.

In some SMD runs using mCEC, the bias forces are so large in the pulling phase that a water molecule will actually auto-ionize. To quantify how frequently this happens we construct water bonding topologies at each timestep by first assigning two closest hydrogens to each oxygen and then assigning the remaining excess proton to its closest oxygen. After we have the bonding topologies, we can monitor the O–H bond lengths of bulk waters that are 6 Å away from the carboxylic oxygens. If a water O–H bond is longer than 1.4 Å, we count it as a broken water. At each timestep, the number of simulations where broken waters are present was counted (Figure 4C). If the mCEC is used in SMD runs, the fraction of broken simulations rises to approximately 0. 5 in the pulling stage, and then remains at this level during equilibration. In comparison, biasing rCEC does not result in any water auto-ionization.

Metadynamics first developed by A. Laio and M. Parrinellos has become increasingly popular as an enhanced sampling method. In metadynamics the bias potential is gradually added in an adaptive fashion, which is significantly different from the case of a stiff moving restraint in SMD simulations. We applied both CECs in a metadynamics simulation of the same Asp system in order to check whether water auto-ionization is an artifact of large bias forces imposed by a short pulling stage in SMD runs. However, water decomposition was still observed when mCEC is biased (Figure S1). Similar water decomposition was observed in QM/MM simulations of human carbonic anhydrase when using mCEC where a water lost a proton to a nearby Histidine.[18, 100] Based on our observations, the resulting “proton-hole” mechanism observed could possibly be an artifact of using mCEC in the reaction coordinate definition. In comparison, metadynamics biasing rCEC and using the same settings otherwise did not show water auto-ionization in the 30 ps run.

3.3. Benchmark of rCEC in umbrella sampling

We then calculated the potential of mean force for the deprotonation of a weak acid biasing our rCEC, and compared it to previous results biasing the exact MS-RMD CEC. Umbrella sampling biasing the MS-RMD CEC has been used extensively, and was shown to reproduce

the experimental pKa when the parameters in the reactive potential are variationally mapped from electronic structure data.[94, 96] Independent umbrella sampling runs biasing rCEC on the same system of water solvated glutamic acid were conducted. The resulting PMF shown in Figure 5A exhibits reasonable agreement with the reference PMF within statistical uncertainties, confirming the stability and usefulness of our new CEC in enhanced sampling.

We further tested rCEC by examining how it performs in other environments. When going to a new system, one could reparametrize rCEC to ensure it reproduces the new charge transfer behavior. However, we show that the MS-RMD CEC position is not strongly dependent on the system itself and thus reparameterization is not necessary. We recalculated the MS-RMD CEC position on the configurations sampled every 2 ps in the original umbrella sampling runs, but using a different MS-RMD model for a glutamic acid developed for ClC-ec1,[101] a Cl^-/H^+ antiporter. As shown in the scatter plot Figure 5B, the resulting CV is nearly identical to the original CV with the largest deviation of only 0.05 Å.

4. Conclusions

In summary, we have discovered valuable spectroscopic properties of the center of excess charge of a hydrated proton, and presented a new CEC (rCEC) definition with improved properties. Since the excess charge vanishes on bulk-like waters, the rCEC spectrum provides a promising computational tool for studying IR properties of the solvated proton in an aqueous environment under room temperature perfectly disentangled from water bands. In turn, the spectral features of different CEC definitions provide insights into each CEC's behavior, which was then validated by analyzing each CEC in various enhanced sampling simulations. In future work, we will apply our rCEC in AIMD simulations and QM/MM simulations of proton transport of more complicated systems. Further decomposing the excess charge according to motions has and will enable us to assign IR signals in more detail without the contamination of water absorption while directly working in condensed phase. Understanding the excess charge from a pure *ab initio* picture may derive from the constrained DFT (CDFT)[102] and multistate DFT (MS-DFT) framework[103, 104] while future efforts are required to accommodate biasing electronic degrees of freedom into Born-Oppenheimer MD if one wants to directly enhance the excess charge transfer via enhanced sampling. One possible solution is parametrizing our rCEC to reproduce the CDFT or MS-DFT charge transfer calculations so that the excess charge transfer can be indirectly biased through our CEC. Future work may also go beyond the center of excess charge, which is essentially the first moment of excess charge distribution. The second moment, which includes the spatial shape of excess charge distribution may also be useful to study recent polarization experiments on acidic solutions.[33, 98]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The personnel in this research were supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH Grant R01-GM053148). The computational resources were provided by the

Research Computing Center at the University of Chicago. The authors thank Greg Voth and Paul Calio for numerous valuable discussions and insights.

REFERENCES

1. Eigen M, Proton Transfer Acid-Base Catalysis + Enzymatic Hydrolysis .I. Elementary Processes. Angew Chem Int Edit 1964, 3 (1), 1–19.
2. Schuster P; Zundel G: Sandorfy C. The Hydrogen bond: recent developments in theory and experiments. North-Holland Pub. Co.; distributor, American Elsevier Pub. Co.: Amsterdam New York, 1976.
3. Markovitch O; Chen H.: Izvekov S; Paesani F; Voth GA; Agmon N, Special pair dance and partner selection: Elementary steps in proton transport in liquid water. J Phys Chem B 2008, 112 (31), 9456–9466. [PubMed: 18630857]
4. Swanson JMJ; Maupin CM; Chen HN; Petersen MK; Xu JC; Wu YJ; Voth GA, Proton solvation and transport in aqueous and biomolecular systems: Insights from computer simulations. J Phys Chem B 2007, 111 (17), 4300–4314. [PubMed: 17429993]
5. Agmon N, The Grotthuss Mechanism. Chem Phys Lett 1995, 244 (5-6), 456–462.
6. Torrie GM; Valleau JP, Non-Physical Sampling Distributions in Monte-Carlo Free-Energy Estimation - Umbrella Sampling. J Comput Phys 1977, 23 (2), 187–199.
7. Bartels C; Karplus M, Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations. J Comput Chem 1997, 18 (12), 1450–1462.
8. Laio A; Parrinello M, Escaping free-energy minima. P Natl Acad Sci USA 2002, 99 (20), 12562–12566.
9. Laio A; Gervasio FL, Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. Rep Prog Phys 2008, 71 (12), 126601.
10. Barducci A; Bonomi M; Parrinello M, Metadynamics. Wires Comput Mol Sci 2011, 1 (5), 826–843.
11. Dama JF; Hocky GM; Sun R; Voth GA, Exploring Valleys without Climbing Every Peak: More Efficient and Forgiving Metabasin Metadynamics via Robust On-the-Fly Bias Domain Restriction. J Chem Theory Comput 2015, 11 (12), 5638–5650. [PubMed: 26587809]
12. Dama JF; Rotskoff G; Parrinello M; Voth GA, Transition-Tempered Metadynamics: Robust, Convergent Metadynamics via On-the-Fly Transition Barrier Estimation. J Chem Theory Comput 2014, 10 (9), 3626–3633. [PubMed: 26588507]
13. Comer J; Gumbart JC; Henin J; Lelievre T; Pohorille A; Chipot C, The Adaptive Biasing Force Method: Everything You Always Wanted To Know but Were Afraid To Ask. J Phys Chem B 2015, 119 (3), 1129–1151. [PubMed: 25247823]
14. Fu HH; Shao XG; Chipot C; Cai WS, Extended Adaptive Biasing Force Algorithm. An On-the-Fly Implementation for Accurate Free-Energy Calculations. J Chem Theory Comput 2016, 12 (8), 3506–3513. [PubMed: 27398726]
15. Pomes R; Roux B, Free energy profiles for H+ conduction along hydrogen-bonded chains of water molecules. Biophys J 1998, 75 (1), 33–40. [PubMed: 9649365]
16. Chakrabarti N; Tajkhorshid E; Roux B; Pomes R, Molecular basis of proton blockage in aquaporins. Structure 2004, 12 (1), 65–74. [PubMed: 14725766]
17. Konig PH; Ghosh N; Hoffmann M; Elstner M; Tajkhorshid E; Frauenheim T; Cui Q, Toward theoretical analyis of long-range proton transfer kinetics in biomolecular pumps. J Phys Chem A 2006, 110 (2), 548–563. [PubMed: 16405327]
18. Riccardi D; Konig P; Prat-Resina X; Yu HB; Elstner M; Frauenheim T; Cui Q, “Proton holes” in long-range proton transfer reactions in solution and enzymes: A theoretical analysis. J Am Chem Soc 2006, 128 (50), 16302–16311. [PubMed: 17165785]
19. Liang RB; Swanson JMJ; Voth GA, Benchmark Study of the SCC-DFTB Approach for a Biomolecular Proton Channel. J Chem Theory Comput 2014, 10 (1), 451–462. [PubMed: 25104919]

20. Liang RB; Swanson JMJ; Madsen JJ; Hong M; DeGrado WF; Voth GA, Acid activation mechanism of the influenza A M2 proton channel. P Natl Acad Sci USA 2016, 113 (45), E6955–E6964.
21. Lee S; Swanson JMJ; Voth GA, Multiscale Simulations Reveal Key Aspects of the Proton Transport Mechanism in the ClC-ec1 Antiporter. Biophys J 2016, 110 (6), 1334–1345. [PubMed: 27028643]
22. Tolle P; Kohler C; Marschall R; Sharifi M; Wark M; Frauenheim T, Proton transport in functionalised additives for PEM fuel cells: contributions from atomistic simulations. Chem Soc Rev 2012, 41 (15), 5143–5159. [PubMed: 22595861]
23. Holub D; Kubar T; Mast T; Elstner M; Gillet N, What accounts for the different functions in photolyases and cryptochromes: a computational study of proton transfers to FAD. Phys Chem Chem Phys 2019, 21 (22), 11956–11966. [PubMed: 31134233]
24. Duster AW; Lin H, Restrained Proton Indicator in Combined Quantum-Mechanics/Molecular-Mechanics Dynamics Simulations of Proton Transfer through a Carbon Nanotube. J Phys Chem B 2017, 121 (36), 8585–8592. [PubMed: 28820594]
25. Woutersen S; Bakker HJ, Ultrafast vibrational and structural dynamics of the proton in liquid water. Phys Rev Lett 2006, 96 (13), 138305. [PubMed: 16712045]
26. Amir W; Gallot G; Hache F; Bratos S; Leicknam JC; Vuilleumier R, Time-resolved observation of the Eigen cation in liquid water. J Chem Phys 2007, 126 (3), 034511. [PubMed: 17249888]
27. Tielrooij KJ; Timmer RLA; Bakker HJ; Bonn M, Structure Dynamics of the Proton in Liquid Water Probed with Terahertz Time-Domain Spectroscopy. Phys Rev Lett 2009, 102 (19), 198303. [PubMed: 19519004]
28. Roberts ST; Petersen PB; Ramasesha K; Tokmakoff A; Ufimtsev IS; Martinez TJ, Observation of a Zundel-like transition state during proton transfer in aqueous hydroxide solutions. P Natl Acad Sci USA 2009, 106 (36), 15154–15159.
29. Roberts ST; Ramasesha K; Petersen PB; Mandal A; Tokmakoff A, Proton Transfer in Concentrated Aqueous Hydroxide Visualized Using Ultrafast Infrared Spectroscopy. J Phys Chem A 2011, 115 (16), 3957–3972. [PubMed: 21314148]
30. Thamer M; De Marco L; Ramasesha K; Mandal A; Tokmakoff A, Ultrafast 2D IR spectroscopy of the excess proton in liquid water. Science 2015, 350 (6256), 78–82. [PubMed: 26430117]
31. Ottosson N; Liu L; Bakker HJ, Vibrational Relaxation of the Aqueous Proton in Acetonitrile: Ultrafast Cluster Cooling and Vibrational Predissociation. J Phys Chem B 2016, 120 (29), 7154–7163. [PubMed: 27333302]
32. Dahms F; Costard R; Pines E; Fingerhut BP; Nibbering ETJ; Elsaesser T, The Hydrated Excess Proton in the Zundel Cation H5O2+: The Role of Ultrafast Solvent Fluctuations. Angew Chem Int Edit 2016, 55 (36), 10600–10605.
33. Carpenter WB; Lewis NHC; Fournier JA; Tokmakoff A, Entropic barriers in the kinetics of aqueous proton transfer. J Chem Phys 2019, 151 (3), 034501. [PubMed: 31325917]
34. Asmis KR; Pivonka NL; Santambrogio G; Brummer M; Kaposta C; Neumark DM; Woste L, Gas-phase infrared spectrum of the protonated water dimer. Science 2003, 299 (5611), 1375–1377. [PubMed: 12574498]
35. Biswas R; Carpenter W; Fournier JA; Voth GA; Tokmakoff A, IR spectral assignments for the hydrated excess proton in liquid water. J Chem Phys 2017, 146 (15), 154507. [PubMed: 28433032]
36. Dahms F; Fingerhut BP; Nibbering ETJ; Pines E; Elsaesser T, Large-amplitude transfer motion of hydrated excess protons mapped by ultrafast 2D IR spectroscopy. Science 2017, 357 (6350), 491–494. [PubMed: 28705988]
37. Fournier JA; Carpenter WB; Lewis NHC; Tokmakoff A, Broadband 2D IR spectroscopy reveals dominant asymmetric H5O2+ proton hydration structures in acid solutions. Nat Chem 2018, 10 (9), 932–937. [PubMed: 30061612]
38. Kulig W; Agmon N, A ‘clusters-in-liquid’ method for calculating infrared spectra identifies the proton-transfer mode in acidic aqueous solutions. Nat Chem 2013, 5 (1), 29–35. [PubMed: 23247174]

39. Xu JQ; Zhang Y; Voth GA, Infrared Spectrum of the Hydrated Proton in Water. J Phys Chem Lett 2011, 2 (2), 81–86. [PubMed: 26295525]
40. Kulig W; Agmon N, Deciphering the infrared spectrum of the protonated water pentamer and the hybrid Eigen-Zundel cation. Phys Chem Chem Phys 2014, 16 (10), 4933–4941. [PubMed: 24477279]
41. Yu Q; Carpenter WB; Lewis NHC; Tokrnakoff A; Bowman JM, High-Level VSCF/VCI Calculations Decode the Vibrational Spectrum of the Aqueous Proton. J Phys Chem B 2019, 123 (33), 7214–7224. [PubMed: 31361141]
42. Daly CA; Streacker LM; Sun YC; Pattenaude SR; Hassanali AA; Petersen PB; Corcelli SA; Ben-Arnotz D, Decomposition of the Experimental Raman and Infrared Spectra of Acidic Water into Proton, Special Pair, and Counterion Contributions. J Phys Chem Lett 2017, 8 (21), 5246–5252. [PubMed: 28976760]
43. Iftimie R; Tuckerman ME, Decomposing total IR spectra of aqueous systems into solute and solvent contributions: A computational approach using maximally localized Wannier orbitals. J Chem Phys 2005, 122 (21), 214508. [PubMed: 15974755]
44. Kim J; Schmitt UW; Gruetzmacher JA; Voth GA; Scherer NE, The vibrational spectrum of the hydrated proton: Comparison of experiment, simulation, and normal mode analysis. J Chem Phys 2002, 116 (2), 737–746.
45. Napoli JA; Marsalek O; Markland TE, Decoding the spectroscopic features and time scales of aqueous proton defects. J Chem Phys 2018, 148 (22), 222833. [PubMed: 29907063]
46. Swanson JMJ; Simons J, Role of Charge Transfer in the Structure and Dynamics of the Hydrated Proton. J Phys Chem B 2009, 113 (15), 5149–5161. [PubMed: 19309128]
47. Lapid H; Agmon N; Petersen MK; Voth GA, A bond-order analysis of the mechanism for hydrated proton mobility in liquid water. J Chem Phys 2005, 122 (1), 014506.
48. Hassanali AA; Cuny J; Verdolino V; Parrinello M, Aqueous solutions: state of the art in ab initio molecular dynamics. Philos T R Soc A 2014, 372 (2011), 20120482.
49. Tuckerman M; Laasonen K; Sprik M; Parrinello M, Ab-Initio Molecular-Dynamics Simulation of the Solvation and Transport of H3o+ and Oh-Ions in Water. J Phys Chem-Us 1995, 99 (16), 5749–5752.
50. Marx D; Tuckerman ME; Hutter J; Parrinello M, The nature of the hydrated excess proton in water. Nature 1999, 397 (6720), 601–604.
51. Marx D, Proton transfer 200 years after von Grotthuss: Insights from ab initio simulations. Chemphyschem 2006, 7 (9), 1848–1870. [PubMed: 16929553]
52. Tse YLS; Knight C; Voth GA, An analysis of hydrated proton diffusion in ab initio molecular dynamics. J Chem Phys 2015, 142 (1), 014104. [PubMed: 25573550]
53. Tuckerman M; Laasonen K; Sprik M; Parrinello M, Ab-Initio Molecular-Dynamics Simulation of the Solvation and Transport of Hydronium and Hydroxyl Ions in Water. J Chem Phys 1995, 103 (1), 150–161.
54. Chandra A; Tuckerman ME; Marx D, Connecting solvation shell structure to proton transport kinetics in hydrogen-bonded networks via population correlation functions. Phys Rev Lett 2007, 99 (14), 145901. [PubMed: 17930688]
55. Berkelbach TC; Lee HS; Tuckerman ME, Concerted Hydrogen-Bond Dynamics in the Transport Mechanism of the Hydrated Proton: A First-Principles Molecular Dynamics Study. Phys Rev Lett 2009, 103 (23), 238302. [PubMed: 20366181]
56. Riera M; Lambros E; Nguyen TT; Gotz AW; Paesani F, Low-order many-body interactions determine the local structure of liquid water. Chem Sci 2019, 10 (35), 8211–8218. [PubMed: 32133122]
57. Medders GR; Babin V; Paesani F, A Critical Assessment of Two-Body and Three-Body Interactions in Water. J Chem Theory Comput 2013, 9 (2), 1103–1114. [PubMed: 26588754]
58. Medders GR; Gotz AW; Morales MA; Bajaj P; Paesani F, On the representation of many-body interactions in water. J Chem Phys 2015, 143 (10), 104102 [PubMed: 26374013]
59. Galib M; Duignan TT; Misteli Y; Baer MD; Schenter GK; Hutter J; Mundyc CJ, Mass density fluctuations in quantum and classical descriptions of liquid water. J Chem Phys 2017, 146 (24), 244501. [PubMed: 28668048]

60. Marsalek O; Markland TE, Quantum Dynamics and Spectroscopy of Ab Initio Liquid Water: The Interplay of Nuclear and Electronic Quantum Effects. J Phys Chem Lett 2017, 8 (7), 1545–1551. [PubMed: 28296422]
61. Ceriotti M; Fang W; Kusalik PG; McKenzie RH; Michaelides A; Morales MA; Markland TE, Nuclear Quantum Effects in Water and Aqueous Systems: Experiment, Theory, and Current Challenges. Chem Rev 2016, 116 (13), 7529–7550. [PubMed: 27049513]
62. Markland TE; Ceriotti M, Nuclear quantum effects enter the mainstream. Nat Rev Chem 2018, 2 (3), 109.
63. Gillan MJ; Alfe D; Michaelides A, Perspective: How good is DFT for water? J Chem Phys 2016, 144 (13), 130901. [PubMed: 27059554]
64. Voth GA, The computer simulation of proton transport in biomolecular systems. Front Biosci-Landmrk 2003, 8, S1384–S1397.
65. Voth GA, Computer simulation of proton solvation and transport in aqueous and biomolecular systems. Accounts Chem Res 2006, 39 (2), 143–150.
66. Lobaugh J; Voth GA, The quantum dynamics of an excess proton in water. J Chem Phys 1996, 104 (5), 2056–2069.
67. Schmitt UW; Voth GA, Multistate empirical valence bond model for proton transport in water. J Phys Chem B 1998, 102 (29), 5547–5551.
68. Cuma M; Schmitt UW; Voth GA, A multistate empirical valence bond model for acid-base chemistry in aqueous solution. Chem Phys 2000, 258 (2-3), 187–199.
69. Cuma M; Schmitt UW; Voth GA, A multistate empirical valence bond model for weak acid dissociation in aqueous solution. J Phys Chem A 2001, 105 (12), 2814–2823.
70. Knight C; Lindberg GE; Voth GA, Multiscale reactive molecular dynamics. J Chem Phys 2012, 137 (22), 22A525.
71. Lee S; Liang RB; Voth GA; Swanson JMJ, Computationally Efficient Multiscale Reactive Molecular Dynamics to Describe Amino Acid Deprotonation in Proteins. J Chem Theory Comput 2016, 12 (2), 879–891. [PubMed: 26734942]
72. Biswas R; Tse YLS; Tokmakoff A; Voth GA, Role of Presolvation and Anharmonicity in Aqueous Phase Hydrated Proton Solvation and Transport. J Phys Chem B 2016, 120 (8), 1793–1804. [PubMed: 26575795]
73. Wu YJ; Chen HN; Wang F; Paesani F; Voth GA, An improved multistate empirical valence bond model for aqueous proton solvation and transport. J Phys Chem B 2008, 112 (2), 467–482. [PubMed: 17999484]
74. Lee S; Mayes HB; Swanson JMJ; Voth GA, The Origin of Coupled Chloride and Proton Transport in a Cl-/H+ Antiporter. J Am Chem Soc 2016, 138 (45), 14923–14930. [PubMed: 27783900]
75. Mayes HB; Lee S; White AD; Voth GA; Swanson JMJ, Multiscale Kinetic Modeling Reveals an Ensemble of Cl-/H+ Exchange Pathways in ClC-ec1 Antiporter. J Am Chem Soc 2018, 140 (5), 1793–1804. [PubMed: 29332400]
76. Liang RB; Li H; Swanson JMJ; Voth GA, Multiscale simulation reveals a multifaceted mechanism of proton permeation through the influenza A M2 proton channel. P Natl Acad Sci USA 2014, 111 (26), 9396–9401.
77. Liang RB; Swanson JMJ; Peng YX; Wikstrom M; Voth GA, Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome c oxidase. P Natl Acad Sci USA 2016, 113 (27), 7420–7425.
78. Liang RB; Swanson JMJ; Wikstrom M; Voth GA, Understanding the essential proton-pumping kinetic gates and decoupling mutations in cytochrome c oxidase. P Natl Acad Sci USA 2017, 114 (23), 5924–5929.
79. Kofinger J; Dellago C, Biasing the center of charge in molecular dynamics simulations with empirical valence bond models: Free energetics of an excess proton in a water droplet. J Phys Chem B 2008, 112 (8), 2349–2356. [PubMed: 18247589]
80. Pauling L, The Nature of the Chemical Bond. Cornell university press Ithaca, NY: 1960; Vol. 260.
81. Madden PA; Impey RW, On the Infrared and Raman-Spectra of Water in the Region 5-250 Cm-1. Chem Phys Lett 1986, 123 (6), 502–506.

82. Chen C; Arntsen C; Voth GA, Development of reactive force fields using ab initio molecular dynamics simulation minimally biased to experimental data. J Chem Phys 2017, 147 (16), 161719. [PubMed: 29096465]

83. Powell MJD, An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput J 1964, 7 (2), 155–162.

84. Tange O, GNU Parallel - The Command-Line Power Tool. ;login: The USENIX Magazine 2015, 36 (1), 42–47.

85. Pezeshki S; Lin H, Adaptive-Partitioning QM/MM for Molecular Dynamics Simulations: 4. Proton Hopping in Bulk Water. J Chem Theory Comput 2015, 11 (6), 2398–2411. [PubMed: 26575540]

86. VandeVondele J; Krack M; Mohamed F; Parrinello M; Chassaing T; Hutter J, QUICKSTEP: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. Comput Phys Commun 2005, 167 (2), 103–128.

87. Lippert G; Hutter J; Parrinello M, A hybrid Gaussian and plane wave density functional scheme. Mol Phys 1997, 92 (3), 477–487.

88. Tribello GA; Bonomi M; Branduardi D; Camilloni C; Bussi G, PLUMED 2: New feathers for an old bird. Comput Phys Commun 2014, 185 (2), 604–613.

89. Hartwigsen C; Goedecker S; Hutter J, Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. Phys Rev B 1998, 58 (7), 3641–3662.

90. Grimme S; Antony J; Ehrlich S; Krieg H, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 2010, 132 (15), 154104. [PubMed: 20423165]

91. White AD; Voth GA, Efficient and Minimal Method to Bias Molecular Simulations with Experimental Data. J Chem Theory Comput 2014, 10 (8), 3023–3030. [PubMed: 26588273]

92. White AD; Knight C; Hocky GM; Voth GA, Communication: Improved ab initio molecular dynamics by minimally biasing with experimental data. J Chem Phys 2017, 146 (4), 041102. [PubMed: 28147531]

93. McCullagh M; Saunders MG; Voth GA, Unraveling the Mystery of ATP Hydrolysis in Actin Filaments. J Am Chem Soc 2014, 136 (37), 13053–13058. [PubMed: 25181471]

94. Nelson JG; Peng YX; Silverstein DW; Swanson JMJ, Multiscale Reactive Molecular Dynamics for Absolute pK(a) Predictions and Amino Acid Deprotonation. J Chem Theory Comput 2014, 10 (7), 2729–2737. [PubMed: 25061442]

95. Yamashita T; Peng YX; Knight C; Voth GA, Computationally Efficient Multiconfigurational Reactive Molecular Dynamics. J Chem Theory Comput 2012, 8 (12), 4863–4875. [PubMed: 25100924]

96. Maupin CM; Wong KF; Soudackov AV; Kim S; Voth GA, A multistate empirical valence bond description of protonatable amino acids. J Phys Chem A 2006, 110 (2), 631–639. [PubMed: 16405335]

97. Headrick JM; Diken EG; Walters RS; Hammer NI; Christie RA; Cui J; Myshakin EM; Duncan MA; Johnson MA; Jordan KD, Spectral signatures of hydrated proton vibrations in water clusters. Science 2005, 308 (5729), 1765–1769. [PubMed: 15961665]

98. Carpenter WB; Fournier JA; Lewis NHC; Tokmakoff A, Picosecond Proton Transfer Kinetics in Water Revealed with Ultrafast IR Spectroscopy. J Phys Chem B 2018, 122 (10), 2792–2802. [PubMed: 29452488]

99. Xu JQ; Izvekov S; Voth GA, Structure and Dynamics of Concentrated Hydrochloric Acid Solutions. J Phys Chem B 2010, 114 (29), 9555–9562. [PubMed: 20593833]

100. Riccardi D; Cui Q, pK(a) analysis for the zinc-bound water in human carbonic anhydrase II: Benchmark for “Multiscale” QM/MM simulations and mechanistic implications. J Phys Chem A 2007, 111 (26), 5703–5711. [PubMed: 17506534]

101. Wang Z; Swanson JMJ; Voth GA, Modulating the Chemical Transport Properties of a Transmembrane Antiporter via Alternative Anion Flux. J Am Chem Soc 2018, 140 (48), 16535–16543. [PubMed: 30421606]

102. Kaduk B; Kowalczyk T; Van Voorhis T, Constrained Density Functional Theory. Chem Rev 2012, 112 (1), 321–370. [PubMed: 22077560]

103. Cembran A; Song LC; Mo YR; Gao JL, Block-Localized Density Functional Theory (BLDFT), Diabatic Coupling, and Their Use in Valence Bond Theory for Representing Reactive Potential Energy Surfaces. J Chem Theory Comput 2009, 5 (10), 2702–2716. [PubMed: 20228960]

104. Ren HS; Provorse MR; Bao P; Qu ZX; Gao JL, Multistate Density Functional Theory for Effective Diabatic Electronic Coupling. J Phys Chem Lett 2016, 7 (12), 2286–2293. [PubMed: 27248004]

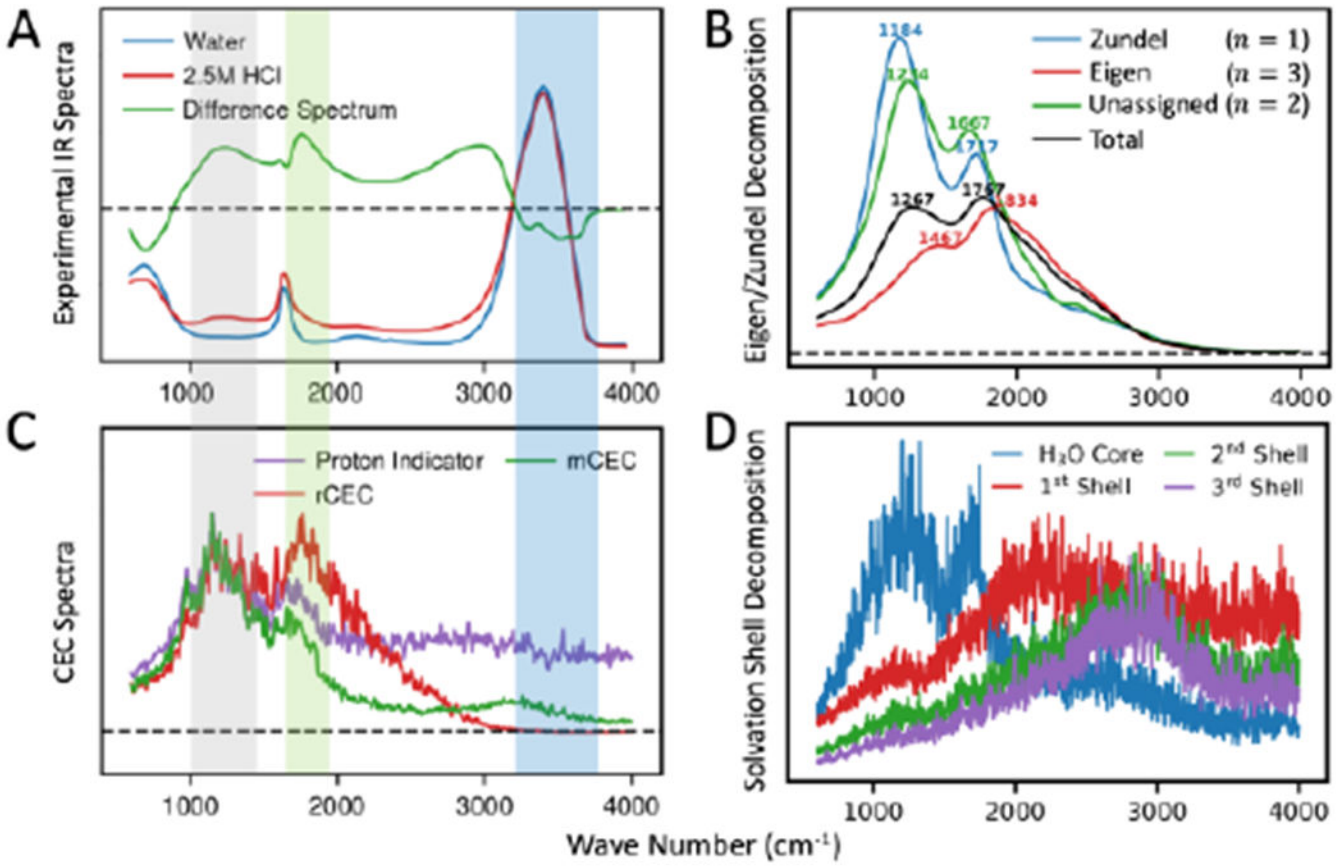


Figure 1.

(A) Experimental IR spectrum of pure water (blue), 2.5M HCl solution (red) and their difference (green). The gray shadow highlights the PTM signal, the green shows the flanking water bending mode and the blue shows pure water stretching. (B) Decomposition of rCEC spectrum by number of unique special pairs (*n*). The position of peaks are denoted as wave numbers. The Zundel is defined as having one special pair ($n = 1$) and the Eigen is defined as having three special pairs ($n = 3$). The remaining case ($n = 2$) was not assigned to Eigen or Zundel but left independent. (C) CEC spectra computed from AIMD simulations. The rCEC spectrum (red) shows good agreement in the PTM and flanking water bending and decays to zero at the same position as experimental difference spectrum. The mCEC spectrum (green) shows non-vanishing signal at pure water stretching band. The proton indicator (purple) absorbs in the full frequency range. (D) Decomposition of rCEC spectrum by solvation shells. The absorption spectrum is blue-shifted with the increasing solvation shell.

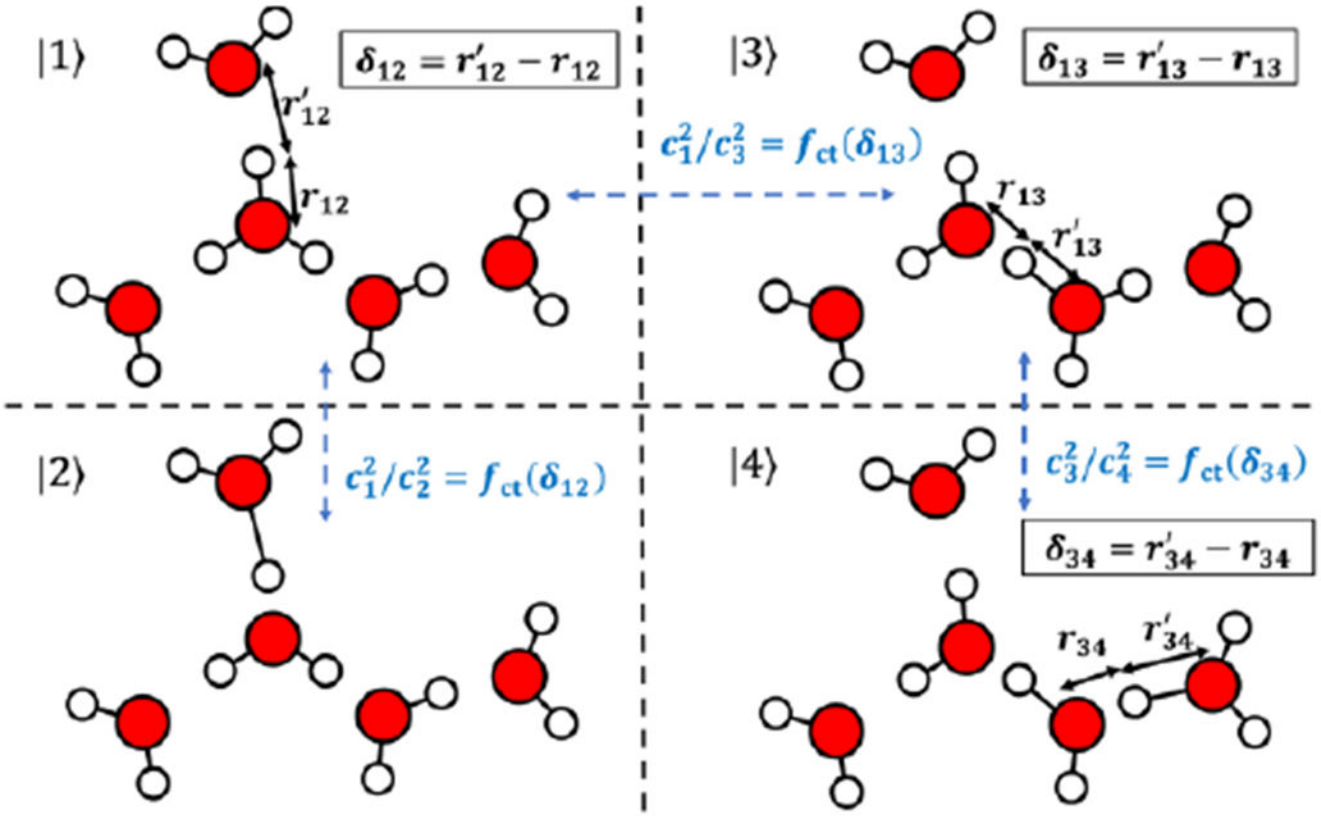


Figure 2.

Illustration of the diabatic states and how the asymmetry coordinate δ and the combining coefficients c_i^2 are calculated for rCEC using $H(H_2O)_{5+}$ as an example. The fifth diabatic state is not shown for clarity.

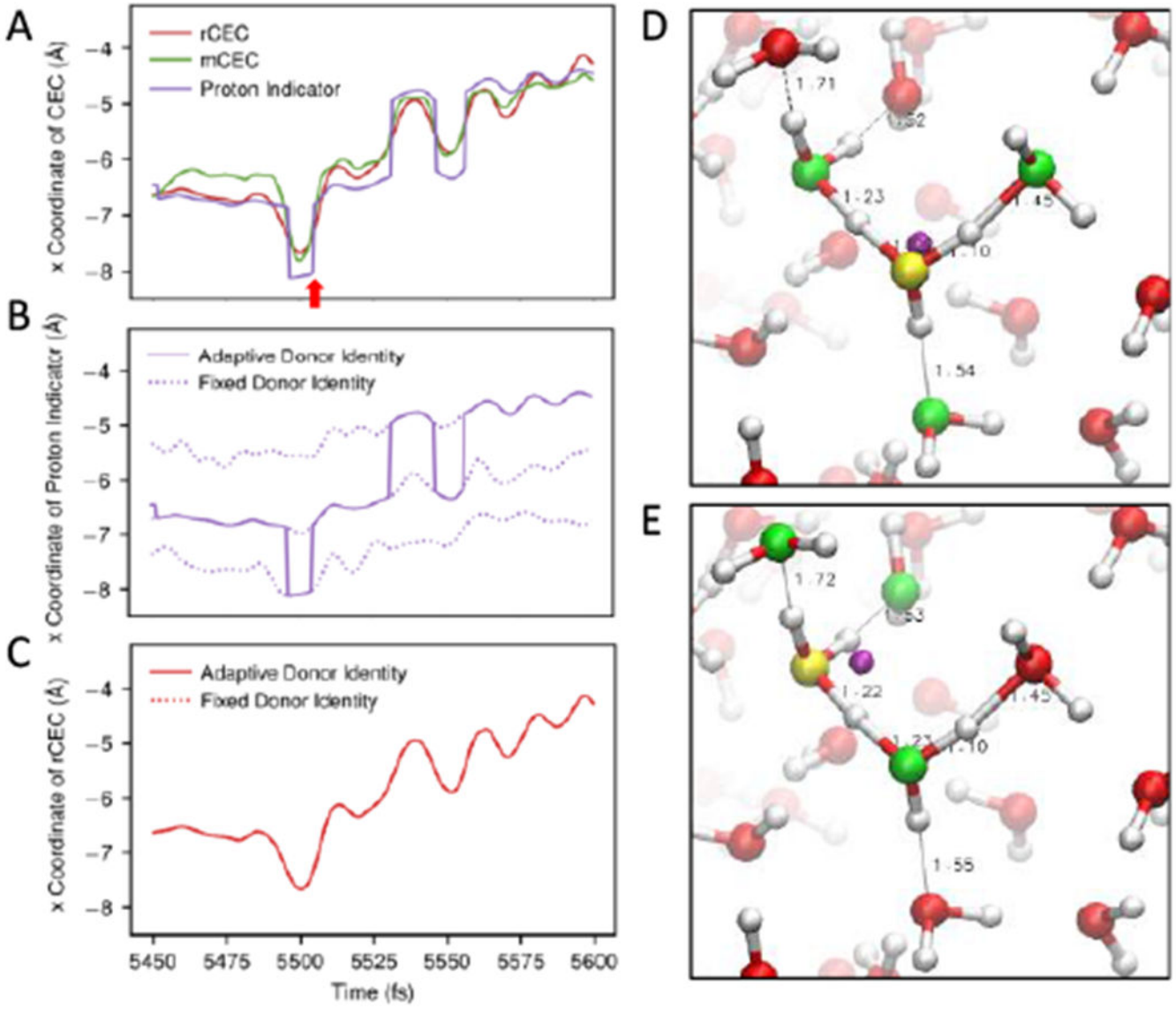


Figure 3.

(A) A representative segment of time series of rCEC (red), mCEC (green) and proton indicator (purple). The red arrow highlights one of the discontinuous jumps of proton indicator and indicates the times of Figure 3D and 3E. (B) The time series of proton indicator in the same time period. The dotted curves show the time evolution of proton indicator with fixed proton donor identity (atom *D* in eq 23) while the solid line shows the one allowing updating the donor identity. (C) The time series of rCEC in the same time period. The dotted and solid curves have the same meaning as Figure 3B. Note that the curves are overlapping perfectly. (D) and (E) Snapshots of AIMD of protonated water from two adjacent timesteps as the red arrow denoted in Figure 3A. The donor oxygen of proton indicator is shown in yellow. The dominant acceptors are shown in green. The position of the proton indicator, shown by a purple ball, jumps over 1 Å in a single step introducing discontinuity.

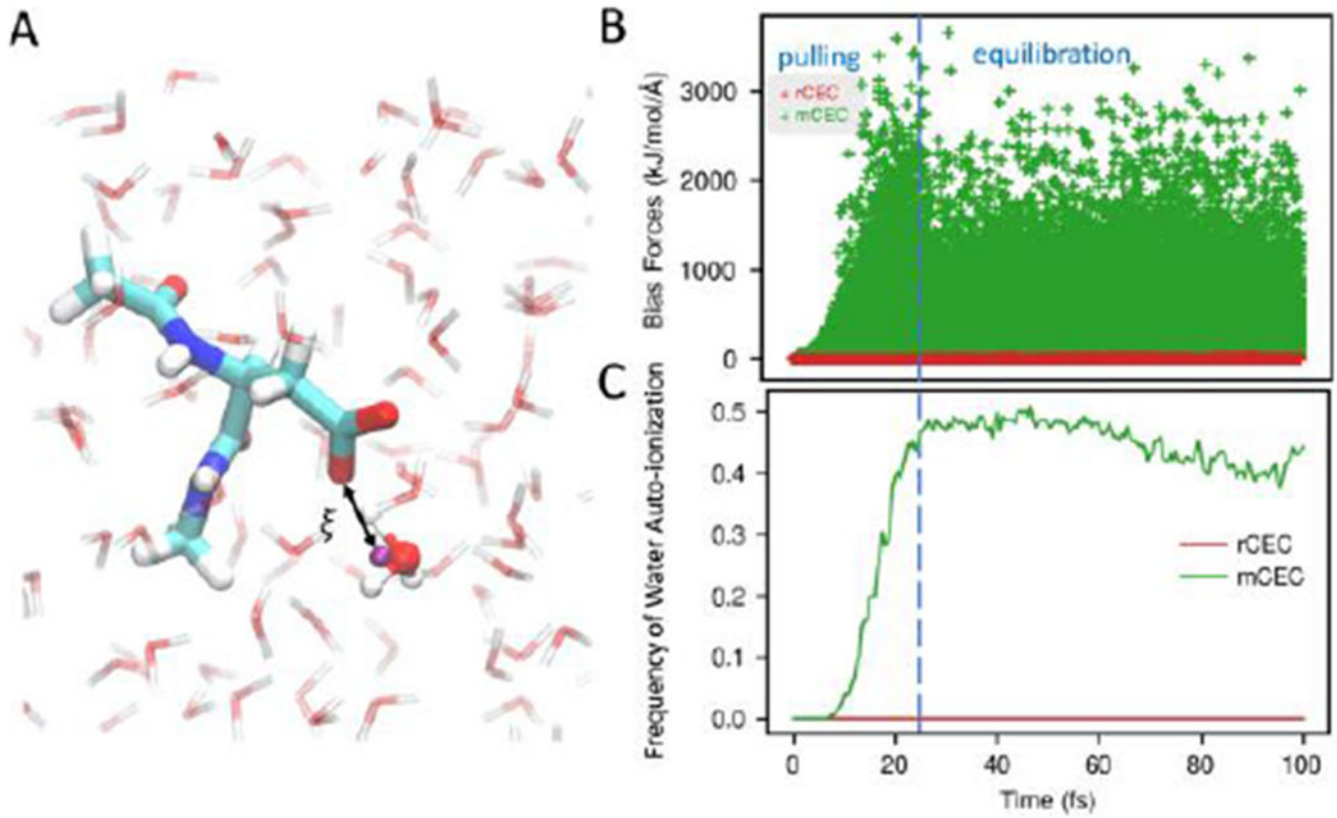


Figure 4.

(A) System setup and collective variable definition of aspartic acid in water. The purple ball represents the CEC. (B) Bias forces acting on bulk water atoms which are beyond 6 Å in all SMD runs. Considerable bias forces exist when biasing mCEC (green) while negligible forces are found in the case of rCEC (red). (C) The ratio of SMD runs where broken water was observed as a function of simulation time. Biasing mCEC results in bulk water decomposition while rCEC is free of this issue.

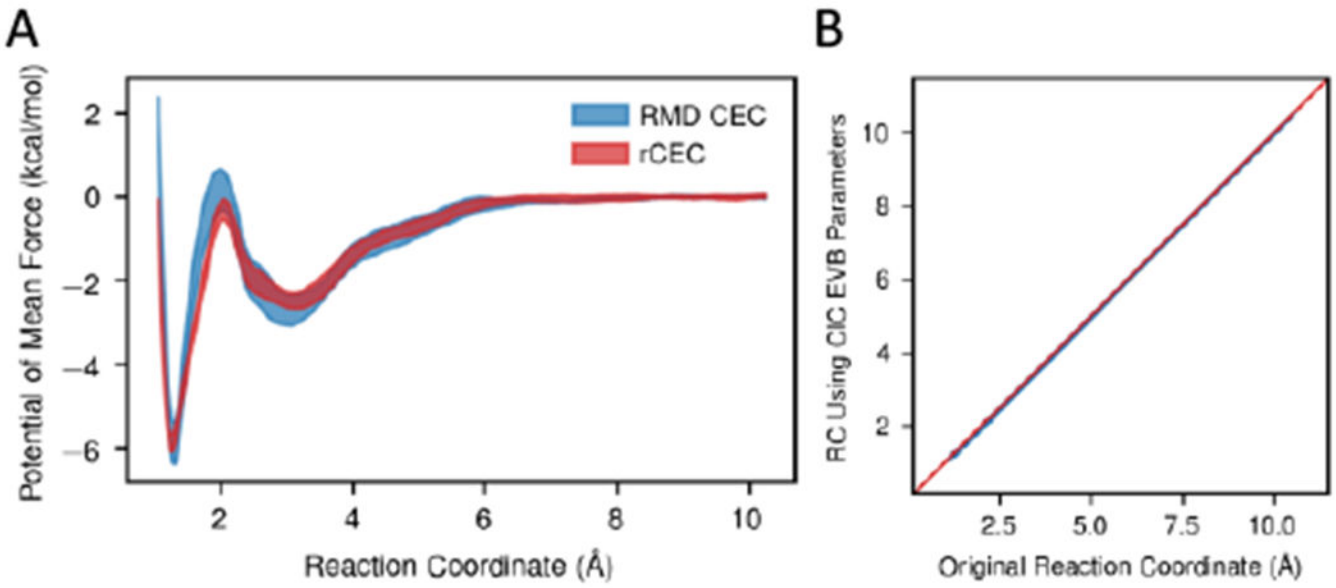


Figure 5.

(A) Deprotonation potential of mean force computed from MS-RMD CEC and rCEC. The two PMFs match within statistical errors. (B) Scatter plot of the reaction coordinate recalculated using a different RMD model with respect to the original reaction coordinate. Data points are shown as blue dots and a red line is for indicating the position of diagonal. The CECs computed from different RMD model match perfectly well.

Table 1.

Parameters of rCEC.

		Initial guess	Final result	Pauling's Bond Order
Hydronium-Water	k (Å^{-1})	4.3937	4.9840	3.8462
	k (Å^{-1})		7.3433	
Amino acid-Water	δ_0 (Å)		0.29316	