Multi-State Reactive Molecular Dynamics Simulations of Proton Diffusion in Water Clusters and in Bulk

Zhen-Hao Xu and Markus Meuwly*

Department of Chemistry, University of Basel, Klingelbergstrasse 80 , CH-4056 Basel, Switzerland.

E-mail: m.meuwly@unibas.ch

Abstract

The molecular mechanics with proton transfer (MMPT) force field is combined with multi state adiabatic reactive molecular dynamics (MS-ARMD) to describe proton transport in the condensed phase. Parametrization for small protonated water clusters based on electronic structure calculations at the MP2/6-311+G(2d,2p) level of theory and refinement by comparing with infrared spectra for protonated water tetramer yields a force field which faithfully describes minimum energy structures of small protonated water clusters. In protonated water clusters up to $(H_2O)_{100}H^+$ the proton hopping rate is around 100 hops/ns. Convergence of such rates is already found for $21 \leq n \leq 31$ and no further speedup in bulk water is found. This indicates that bulk-like behaviour requires solvation of a Zundel motif by $\sim$ 25 water molecules which corresponds to the second solvation sphere. For smaller cluster sizes the number of available states, i.e. the number of proton acceptors, is too small and slows down proton transfer

*To whom correspondence should be addressed

rates. The cluster simulations confirm that the excess proton is typically located on the surface. The free energy surface as a function of the weights of the two lowest states and a configurational parameter suggest that the "special pair" plays a central role in rapid proton transport. The barriers between this minimum energy structure and the Zundel and Eigen minima are sufficiently low ($\sim$ 1 kcal/mol, **consistent with recent experiments and** commensurate with a hopping rate of $\sim$ 100/ns or 1 every 10 ps) which lead to a highly dynamical environment. These findings are also consistent with recent experiments which find that Zundel-type hydration geometries are prevalent in bulk water.

1 Introduction

Proton transfer (PT) is ubiquitous in physical, chemical and biological processes, including PT in liquid water,[1–3] enzymatic catalysis[4] or protein-assisted proton transport in membranes.[5,6] One of the greatest challenges in understanding and characterizing proton translocation is to link kinetic or spectroscopic information with structural changes and the underlying energetics. In many cases, experiments provide thermally and/or structurally averaged data which is difficult to relate with atomistic or molecular aspects involving the motion of the atoms on the relevant time scales, although exceptions to this exist.[7] Spectroscopically, PT has been characterized for a range of systems.[3,8–14]

Proton transfer in bulk water and in clusters involves two processes: oscillatory proton transfer within a Zundel (Z) motif $H_5O_2^+$ and transfer from one Zundel motif to another one ($H_5O_2^+ \cdots H_2O \leftrightarrow H_2O \cdots H_5O_2^+$) which is the so-called Grotthuss mechanism.[2] Gas phase spectroscopic work for small molecular frameworks has found an empirical relationship between the maximum of the infrared (IR) absorption band and the barrier height of PT reactions in combined investigations of computational simulations and experiments.[12,14] Such models can, however, not be applied to the situation in the bulk, such as proton transfer in water, because the Grotthuss mechanism involves transfer of charge through space.

In aqueous systems, the network of hydrogen bonds is a decisive factor in controlling charge transport.[15] Specifically the dynamics in the second solvation shell around the protonation site has been found to be essential for spatial proton transport.[16,17] Grotthuss transport itself is considered a picosecond process[1,18] although direct experimental measurement of this time scale is challenging.[19] The picosecond time scale makes ab initio molecular dynamics (AIMD) an attractive alternative.[1,20,21] Still, such approaches are usually limited to relatively short time scales (tens of picoseconds)[22] and small systems due to the appreciable computational cost involved. Nevertheless, depending on the level of theory, such *ab initio* MD simulations

can provide important direct and molecular level insight into the energetics, dynamics and mechanisms underlying PT in solution.

To study proton translocation in (extended) condensed phase systems and on longer (nanosecond and longer) time scales, there has also been considerable interest in developing empirical methods capable of simulating Grotthuss proton transport. One such method is based on the empirical valence bond (EVB) method.[23,24] Following this, models such as two-state EVB (TS-EVB)[25] and multi state EVB (MS-EVB)[26–31] have been developed. Within this formalism, the electronic wave function is represented as a linear combination of empirical valence bond states

$$|\Psi >= \sum_{i=1}^{N} c_i \cdot |\psi_i > \tag{1}$$

One of the conceptual challenges with EVB-based methods is the need to parametrize the off-diagonal elements which describe the coupling between the states. In the simplest case, the coupling is a scalar but quite often they depend on one or a few selected, geometrical coordinates ("reaction coordinates") which need to be intuitively chosen or motivated a posteriori. Another potentially challenging aspect in concrete applications is the definition of the valence bond states themselves.[32]

Another method for studying proton transfer in molecular systems is Molecular Mechanics with Proton Transfer (MMPT).[33] In this approach, an accurate multi-dimensional potential energy surface (PES) for the proton transfer energetics between an acceptor and a donor atom is parametrized from *ab initio* reference data for generic motifs: symmetric single minimum (SSM; reference system is $O_2H_5^+$), symmetric double minimum (SDM; $N_2H_7^+$) and asymmetric double minimum (ADM; $NH_4^+OH_2$). The parametrization is based on Morse functions which allow hydrogen/proton transfer within a donor/acceptor motif.

Here, MMPT is combined with multi state adiabatic reactive MD (MS-ARMD)[34–36] to follow

proton transport in the condensed phase. The method is referred to as MS-MMPT in the following. The approach taken in the present work is bottom-up in that first the energetics of the Zundel ion is treated within the MMPT framework. Next, the possibility for proton transfer involving a hydrated Zundel ion is provided by using MS-ARMD. Finally, the energy function is reparametrized by comparing with experimental data from small protonated water clusters. This model is then applied to larger clusters and proton transport in the bulk.

2 Methods and Development

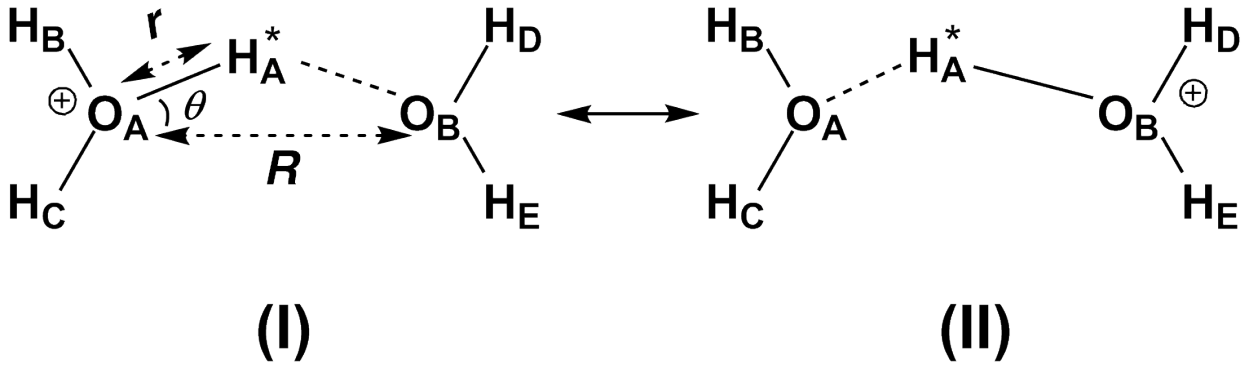


Figure 1: The two resonance structures I and II of protonated water dimer: $H_3O^+ \cdots H_2O$ and $H_2O \cdots H_3O^+$ (important internal coordinates are labeled).

2.1 MMPT force field for $H_5O_2^+$

Figure 1 shows a typical structure of a Zundel ion ($H_5O_2^+$) together with the labeled atoms and important internal coordinates: r, R and θ. Transfer of a proton/hydrogen atom between a donor (D) and an acceptor (A) is characterized by coordinates r ($O_A - H_A^*$) and R ($O_A - O_B$), see Figure 1. The angular displacement of the transferring H_A^* away from the linear path is θ, which is the $H_A^* - O_A - O_B$ angle. **It is advantageous to introduce the unitless progression coordinate** $\rho = (r \cdot cos\theta - r_{\min})/(R - 2r_{\min})$ **[33,37] with** $r_{\min} = 0.8$ **Å which maps** ρ **to the interval** $[1, 0]$**. This choice for** $r_{\min}$ **guarantees that the entire**

attractive region is included in the PES scan for all configurations (R, θ)**.** The MMPT potential is

$$V_{\mathrm{MMPT}}(\rho, R, d) = V_0(\rho, R) + k \cdot d^2, \tag{2}$$

where $V_0(\rho, R)$ is the isotropic part and $k \cdot d^2$ ($d = r \cdot sin\theta$) is a harmonic approximation to the bending mode of $\mathrm{H_A^*}$. The radial dependence of $V_0(\rho, R)$ is represented as a superposition of Morse potentials. Such parametrizations have been used in previous work[12,14,38] to investigate spectroscopic features of PT in small molecular systems.

In the standard MMPT force field, the total energy for one of the two possible resonance states I and II of the Zundel ion, see Figure 1, is the sum of all bonded terms (indicated by solid lines in Figure 1) plus the energy of the $\mathrm{O_A - H_A^* - O_B}$ motif which is $V_{\mathrm{PT}}(\rho, R, d)$. Because the PT motif is explicitly parametrized, no independent $\mathrm{H_3O^+}$ ion exists in the present framework. Within MMPT the energetics involving the transferring $\mathrm{H_A^*}$ is described by energies fitted to reference *ab initio* energies whereas the remaining (bonded and non-bonded terms) are those from an empirical force field, such as CHARMM.[39,40] Upon transfer of $\mathrm{H_A^*}$ (e.g. during a dynamics simulation), particular terms in the two resonance structures are smoothly turned off (e.g. the angle $\mathrm{H_B O_A H_A^*}$) and turned on (e.g. the angle $\mathrm{H_A^* O_B H_D}$). This is achieved by using a geometry-dependent switching function such as

$$\gamma(R, r, \theta) = \frac{1}{2}\big[\tanh[2R \cdot (r \cdot \cos\theta - 0.5R)] + 1\big] \tag{3}$$

All terms that are removed in one resonance structure and generated in another one are treated in this same manner.

In addition to switching on and off particular terms it is also necessary to modify certain internal (bonded) terms between the two resonance structures. Considering resonance structure (I), the bond $\mathrm{O_A H_B}$ is that of an $\mathrm{H_3O^+}$ ion which eventually (after transfer of $\mathrm{H_A^*}$)

becomes a water OH-bond. In other words, the equilibrium bond length and force constant for this bond have to switch when a transition from resonance structure I to II occurs. This is necessary because during the course of an entire simulation the moiety $H_BO_AH_C$ has to be able to become a "regular solvent water" again. Such changes are also described by the switching function γ by representing, for example, the O_AH_B bond energy as

$$E_{\text{bond},0}(l) = (1-\gamma) \cdot k_{\text{b}}^{\text{H}_3\text{O}^+} \cdot (l - l_{\text{eq}}^{\text{H}_3\text{O}^+})^2 + \gamma \cdot k_{\text{b}}^{\text{H}_2\text{O}} \cdot (l - l_{\text{eq}}^{\text{H}_2\text{O}})^2 \tag{4}$$

where l is the bond length of $O_A - H_B$ and $(k_{\text{b}}^{\text{H}_3\text{O}^+}, l_{\text{eq}}^{\text{H}_3\text{O}^+})$ and $(k_{\text{b}}^{\text{H}_2\text{O}}, l_{\text{eq}}^{\text{H}_2\text{O}})$ are the force field parameters for an O–H bond in a hydronium ion and in water, respectively. In order to bring Eq. 4 into a conventional force field form, i.e. $E_{\text{bond}} = k_{\text{eff}} \cdot (l - l_{\text{eq}}^{\text{eff}})^2$, the expression

$$E_{\text{bond}}(l) = E_{\text{bond},0}(l) - E_{\text{sup}}(\gamma) \tag{5}$$

is used with[38]

$$E_{\text{sup}}(\gamma) = \frac{(l_{\text{eq}}^{\text{H}_3\text{O}^+} - l_{\text{eq}}^{\text{H}_2\text{O}})^2}{\frac{1}{(1-\gamma) \cdot k_{\text{b}}^{\text{H}_3\text{O}^+}} + \frac{1}{\gamma \cdot k_{\text{b}}^{\text{H}_2\text{O}}}}. \tag{6}$$

These modifications are applied to bond and angle terms (see Table S1). Such mixed energy terms have been employed previously for spectroscopic studies.[12,38]

To correctly describe the non-planar structure of $H_5O_2^+$, torsional terms $E_{\text{torsion}}(\Psi_{i,j})$ were also introduced for each $[H_iO_AO_BH_j]$ dihedral angle ($i \in \{B, C\}$ and $j \in \{D, E\}$) (see Figure 1), written as

$$E_{\text{torsion}}(\Psi_{i,j}) = k_d \cdot (1 + \cos(\Psi_{i,j})) \tag{7}$$

where k_d is the force constant and $\Psi_{i,j}$ is the dihedral angle.

The resonance structures (see Figure 1) also influence the non-bonded parameters of atoms

in an $H_5O_2^+$ ion depending on the position of the excess proton along the O_AO_B bond. For this, a fluctuating charge model within the $H_5O_2^+$ ion is used whereby the transferring H^* has a fixed partial charge, q_{H^*} and for each non-H^* atom the charge is $q_H = (1-\gamma)\cdot q_O^{H_3O^+} + \gamma \cdot q_O^{H_2O}$ where $q_O^{H_3O^+}$ is the charge of a hydronium oxygen and $q_O^{H_2O}$ is the charge of a water oxygen with γ from Equation 3. The van der Waals parameters (ϵ and $r_{\min}$) for both oxygen atoms within $H_5O_2^+$ were treated in the same fashion while no Lenard Jones potential was applied on H atoms.

Combining all energy terms, the total energy of the $H_5O_2^+$ motif is

$$\begin{aligned} V_{H_5O_2^+}^{\text{intra}} &= V_{\text{PT}}(\rho, R, d) + \sum_{i,j} E_{\text{bond}}(O_iH_j) \\ &+ \sum_{i,j,k} E_{\text{angle}}(O_iH_jO_k) + \sum_{i,j} E_{\text{torsion}}(H_iO_AO_BH_j) \end{aligned} \tag{8}$$

where $\sum E_{\text{bond}}(O_iH_j)$, $\sum E_{\text{angle}}(O_iH_jO_k)$ and $\sum E_{\text{torsion}}(H_iO_AO_BH_j)$ include all atom pairs as indicated in Figure 1, except $O_A - H_A^*$ and $O_B - H_A^*$ bonds which are described by the MMPT potential.

2.2 Multi-state MMPT

Up to this point, transport of H_A^* (see Figure 1) is only possible within a predefined donor–acceptor motif. In order to allow diffusion of charge in a condensed-phase context, additional states have to be introduced. This is done following the general concept of multi state reactive MD (MS-ARMD) which determines the energy and weights of relevant bonding patterns ("states") based on an empirical energy function[34,35,41] which is MMPT in the present case. For a system of n water molecules and an excess proton, a "motif" (or state) in the present context is defined as an $H_5O_2^+$ ion and $(n-2)$ water molecules, see Figure 2.

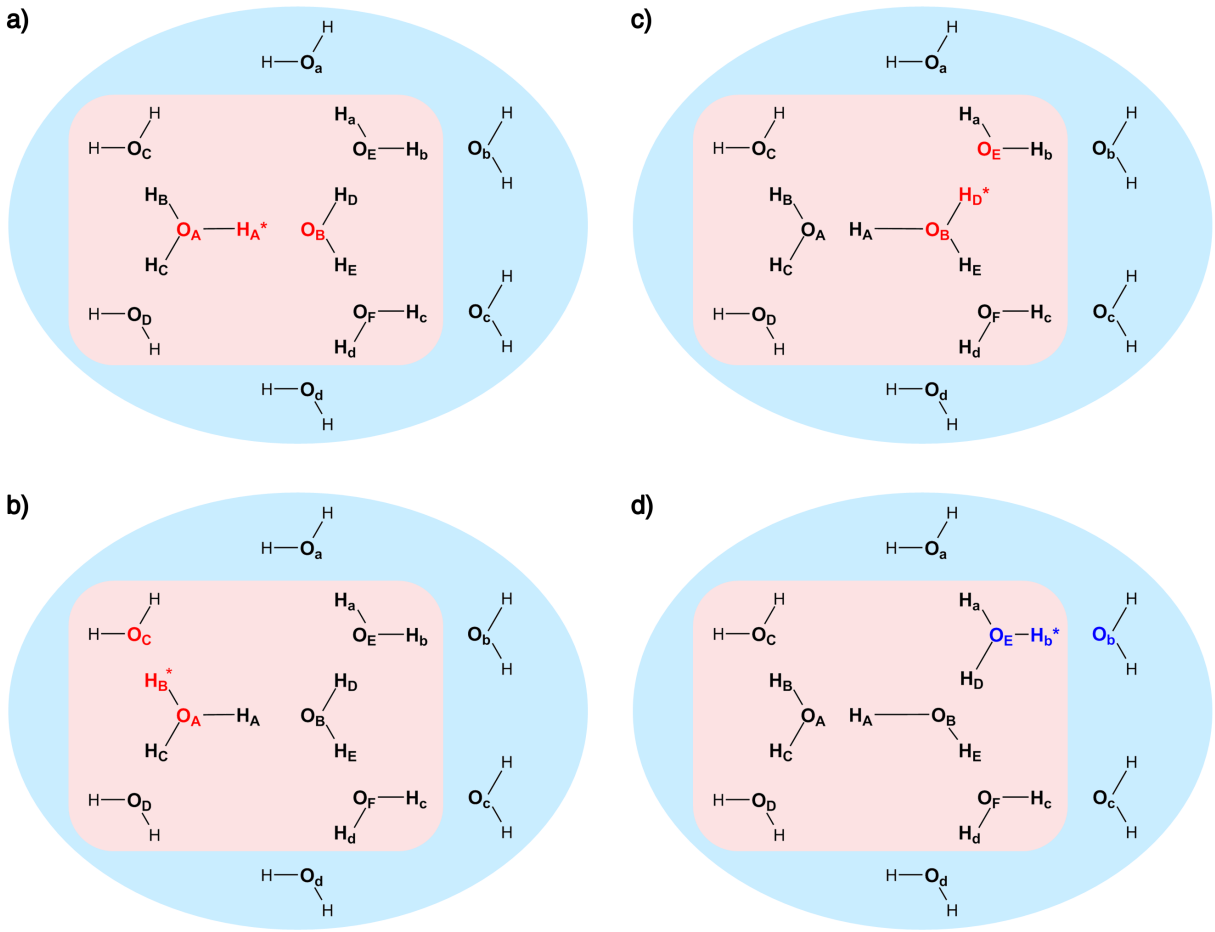


Figure 2: **Schematic representation of** typical states with PT motifs in the MS-MMPT scheme. They include a) the primary state with a 'seed' motif of $O_A - H^*_A - O_B$ and candidate states with motifs b) $O_A - H^*_B - O_C$, c) $O_B - H^*_D - O_E$ and d) $O_E - H^*_b - O_b$. Atoms O_C to O_F are potential H^* acceptors from the first solvation shell (within the red background) and atoms O_a to O_d are those from the second shell to give rise for a second-shell motif such as $O_E - H^*_b - O_b$.

For spatial transport of an excess charge, possible $H_5O_2^+$ motifs are determined for a given spatial arrangement $\mathbf{x}$ of the atoms where $\mathbf{x}$ contains the Cartesian coordinates of all atoms. This is done based on geometrical criteria as described below. The total potential energy of one such state j is

$$V_j(\mathbf{x}) = V^{\text{intra}}_{H_5O_2^+} + V^{\text{intra}}_{\text{solv}} + V^{\text{inter}}_{H_5O_2^+-\text{solv}} + V^{\text{inter}}_{\text{solv}-\text{solv}} \tag{9}$$

where $V^{\text{intra}}_{\text{H}_5\text{O}_2^+}$ is given by Eq. 8, $V^{\text{intra}}_{\text{solv}}$ includes all intramolecular (bonded) energies of solvent molecules and $V^{\text{inter}}_{\text{H}_5\text{O}_2^+-\text{solv}}$ and $V^{\text{inter}}_{\text{solv}-\text{solv}}$ refer to the intermolecular interactions of H_5O_2^+–solvent and solvent–solvent pairs.

Following MS-ARMD, the total potential energy for a given configuration $\mathbf{x}$ is then a linear combination of the energies of the m candidate states

$$V_{\text{tot}}(\mathbf{x}) = \sum_{j=1}^{m} w_j(\mathbf{x}) \cdot V_j(\mathbf{x}), \tag{10}$$

where

$$\begin{aligned} w_j(\mathbf{x}) &= \frac{w_j^0(\mathbf{x})}{\sum_{j=1}^{m} w_j^0(\mathbf{x})} \\ w_j^0(\mathbf{x}) &= \exp\left(-\frac{V_{[\text{H}_5\text{O}_2^+]_j}(\mathbf{x}) - V_{\text{H}_5\text{O}_2^+,\text{min}}(\mathbf{x})}{\Delta E}\right). \end{aligned} \tag{11}$$

In Eq. 10, $w_j(\mathbf{x})$ is the normalized weight[35,42] which has also been used for mixing double many-body expansion PESs.[43] However, other mixing functions are possible such as a tangent hyperbolic[34,41] and thus, no physical significance should be attributed to the mixing function. From a practical perspective, the exponential dependence on the energy to mix the states is akin to a softmax function, as used in neural networks, to ensure normalized weights. Following the concept of MS-ARMD, using exponential weights ensures that the total energy is always larger or equal to the energy of the motif with the lowest energy, i.e. it weighs the most probable state most. The weight $w_j^0(\mathbf{x})$ is determined from the total energy of the jth H_5O_2^+ motif and $V_{\text{H}_5\text{O}_2^+,\text{min}}(\mathbf{x})$ is the H_5O_2^+ motif with the lowest MMPT energy among all candidate states. ΔE is a model parameter which plays a comparable role as the switching time t_s in time-based ARMD.[34,41] It controls the contribution of each selected state to $V_{\text{tot}}(\mathbf{x})$, and can be determined either by fitting it to *ab initio* data[44,45] or by comparing observables computed from an explicit simulation with reference experimental data.[46]

The gradient of the total potential energy is then readily available from

$$\begin{aligned}\nabla V_{\mathrm{tot}}(\mathbf{x}) &= \nabla\Big[\sum_{j=1}^{m} w_j(\mathbf{x}) \cdot V_j(\mathbf{x})\Big] \\ &= \frac{1}{\Delta E} \cdot \frac{1}{\sum_{j=1}^{m} w_j^0(\mathbf{x})} \cdot \sum_{j=1}^{m} \Big\{ w_j^0(\mathbf{x}) \cdot \big[\Delta E \cdot \nabla V_j(\mathbf{x}) + (\,\boldsymbol{V}_{\mathrm{tot}}(\mathbf{x}) - V_j(\mathbf{x})) \cdot \nabla V_{[\mathrm{H_5O_2^+}]_j}(\mathbf{x})\big] \Big\}\end{aligned} \tag{12}$$

which is required for MD simulations.

From Equation 11 it is understood that the state with the lowest $H_5O_2^+$ energy among all candidate states is the state which contributes most to the total potential energy. Because the force field is reactive, it is possible for other low energy states, such as shown in Figure 2, to become the minimum energy state during dynamics simulations which eventually corresponds to a surface crossing between the current and a new $H_5O_2^+$ motif. This corresponds to Grotthuss proton transport.

Candidate states are defined by geometrical criteria which include i) the donor–acceptor ($D-A$) distance to be < 4 Å and ii) the $D-H^*-A$ angle to be $> 90°$. Starting from the $H_5O_2^+$ motif with the lowest energy (the "primary state" or "seed"), D and A are potential H^*-donors for candidate states. Next, two spherical regions with a radius of 4 Å around D and A are considered, which define the first solvation shell. Within the first solvation shell, all O-atoms (other than D or A) are potential H^*-acceptors. Based on these acceptors, new PT motifs and candidate states are constructed. In the same fashion, the second solvation shell motifs and states are constructed, again using a radius of 4 Å. All simulations described in the following use second shell water molecules unless otherwise specified.

As a concrete example the situation in Figure 2 is considered. Here, the primary state involves the $O_A - H_A^* - O_B$ motif (Figure 2a). Based on this motif the first- and second-

solvation shells are built, see red and blue shaded areas. Within the first solvation shell, four additional water molecules, with oxygens O_C to O_F, are found. From this, new PT motifs and corresponding states are generated. Figures 2b and c show examples for such first-shell motifs including $O_A - H^*_B - O_C$ and $O_B - H^*_D - O_E$ for which (see Figure 2c) the H_A atom changes its connectivity from $O_A - H_A$ to $O_B - H_A$. In Figure 2d a second-shell motif is displayed. The bond network is further changed from $O_B - H_D$ to $O_E - H_D$ (compared to Figure 2c) and the motif is $O_E - H^*_b - O_b$. In MS-MMPT simulations, the list of PT donors/acceptors and the corresponding states is updated at each MD time step. A flowchart of the protocol is given in Figure S1.

For computational efficiency it is advantageous to first determine $V^{\text{intra}}_{H_5O_2^+}$ for all candidate states m as this is an inexpensive calculation. If the weight of state i is above a certain threshold (typically $w^i \geq 10^{-11}$; the effect of this choice on the results was tested) this state is regarded as a potential candidate state and the full potential energy is computed according to Eq. 9. Therefore, the total potential energy for a given nuclear configuration is $V_{\text{tot}}(\mathbf{x}) = \sum_{i=1}^{m_{\text{eff}}} w_j(\mathbf{x}) \cdot V_i(\mathbf{x})$ where m_{eff} is the number of states retained according to the criterion for a minimal weight.

2.3 MS-MMPT Simulations and Analysis

Using MS-MMPT, all MD simulations were performed for water clusters ($[H_2O]_nH^+$, $n = 6, 10, 21, 31, 50$ and 100) in the gas phase and water bulk with one excess proton, using CHARMM.[40] To avoid decomposition of the clusters all atoms are constrained with a weak harmonic constraint ($k = 0.001$ kcal·mol^{-1}·Å^{-2}) to the center of mass of the cluster. Starting from the energy optimized structure, heating and equilibration (NVT, 300 K) simulations were run for 100000 MD steps with a time-step of $\Delta t = 0.1$ fs because bonds involving H-atoms are flexible. Such a time step is also used in *ab initio* MD simulations.[22] Then,

1 ns production simulations were carried out in the microcanonical ensemble (NVE) for 10 independent trajectories. For the bulk systems, 16 MD trajectories, each 500 ps in length, were run in a similar fashion (without constraint). The solvent water molecules were modeled by a flexible SPC/fw water model[47] but with the MP2/6-311+G(2d,2p) equilibrium ($l_{\mathrm{eq}} = 0.966$ Å, $\theta_{\mathrm{eq}} = 106.6°$) bond length. Snapshots for analysis were collected every 5 fs.

For determining the proton hopping rates, the time-course accumulation function, $h(t)$, is recorded[29]

$$h(t) = h(t - \delta t) + \delta h(t). \tag{13}$$

Here, t is the simulation time and δt is the lag time $\delta t = 5$ fs between the frames. Initially, $h(t = 0) = 0$, and for each MD frame $\delta h(t)$ assumes one of three values: 0, 1 and -1. For $\delta h(t) = 0$ the proton does not hop whereas for $\delta h(t) = +1$ the identity of the donor atom changes and for $\delta h(t) = -1$ the proton transfers back to the donor from where it came.

3 Results

3.1 Parametrization of the MS-MMPT Force Field

The parametrization of the MS-MMPT force field was carried out in two steps. First, parametrization with respect to the interaction energies (called MS-MMPT-MP2) of a conformational ensemble sampled from finite-temperature MD simulations of small clusters ($H_5O_2^+$ to $H_9O_4^+$) and secondly a refinement based on comparison with experimental infrared spectra (MS-MMPT-IR) which directly report on the curvature of the PES. Such an approach combines parametrization with respect to computed and experimentally measured reference data which has also been found to be beneficial for other parametrized PESs.[48]

The parameters for the 3-dimensional reactive MMPT potential (Eq. 2) were determined by fitting to *ab initio* calculations at the MP2/6-311G++(2d,2p) level.[33,49] The fitting quality is shown in the inset of Figure 3a ($R^2 = 0.999$) and the parameters are given in Table S2. Next, the total energy for the entire $H_5O_2^+$ ion was fitted to reference data at the same level of theory using two resonance structures as shown in Figure 1. For the water molecule the non-bonded parameters of the SPC model[50] were used. The progress of this fit is shown in Figure S2 and the resulting parameters are summarized in Table S3. The quality ($R^2 = 0.88$) of this fit from the room temperature simulations is shown in Figure 3a.

Table 1: Geometric values for optimized $[H_2O]_nH^+$ conformations in Figure 4. Data were calculated using the MS-MMPT-MP2 and MS-MMPT-IR force fields and at the MP2/6-311++G(2d,2p) level of theory. The RMSD between the reference (MP2/6-311++G(2d,2p)) and MS-MMPT-MP2 or MS-MMPT-IR is 0.03 Å and 0.02 Å for bond lengths, and 3.1° and 4.4°, respectively.

$[H_2O]_nH^+$	coordinate	MS-MMPT-MP2	MS-MMPT-IR	MP2
$n = 2$	r, Å	1.179	1.198	1.196
	R, Å	2.395	2.396	2.389
	ϕ	117.0°	117.6°	114.4°
$n = 3$	r, Å	1.019	1.040	1.029
	R, Å	2.514	2.514	2.499
	ϕ	110.4°	110.4°	114.1°
$n = 4$	r, Å	1.008	1.028	1.005
	R, Å	2.536	2.536	2.566
	ϕ	108.4°	108.1°	112.0°
$n = 4$, ring	r, Å	1.025	1.047	1.038
	R_1, Å	2.485	2.489	2.488
	R_2, Å	2.830	2.825	2.851
	ϕ	108.4°	99.9°	106.2°
$n = 6$	r_1, Å	1.233	1.216	1.195
	R_1, Å	2.464	2.433	2.389
	r_2, Å	0.986	0.981	0.983
	R_2, Å	2.611	2.669	2.678

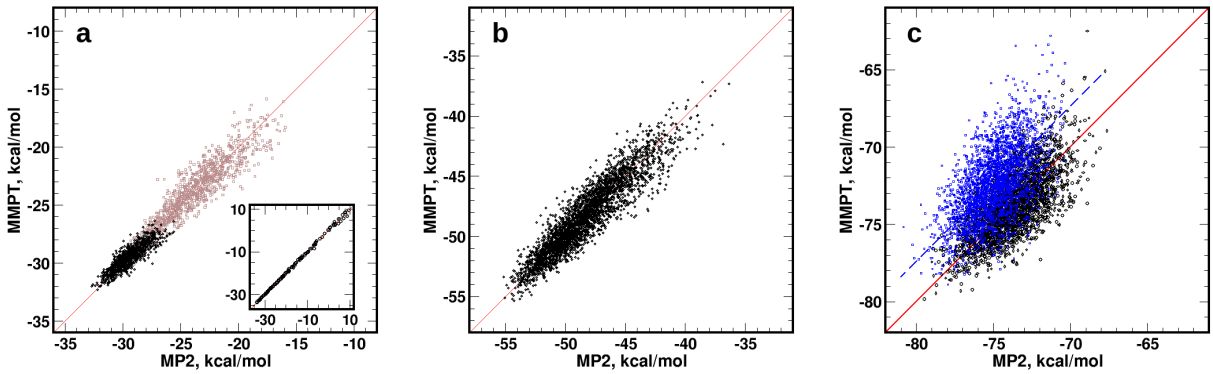


Figure 3: The interaction energies of arbitrary MD frames from MMPT simulations of a) $H_5O_2^+$, b) $[H_2O]_3H^+$ and c) $[H_2O]_4H^+$, compared to MP2 calculations. The fitting quality of the MMPT potential ($V_0(R, \rho)$, see Eq. 2) to the *ab initio* scans (MP2, 253) is shown in the inset in panel a (correlation coefficient $R^2 = 0.999$). In panel a, MD frames were obtained from simulations at 300 K (black, $R^2 = 0.88$) and 750 K (brown, $R^2 = 0.92$). For panel b, MD frames were obtained from simulations at 300 K (black, $R^2 = 0.93$); For panel c, MD frames were obtained from simulations at 300 K using the MS-MMPT-MP2 (black, $R^2 = 0.69$) and MS-MMPT-IR force field (blue, $R^2 = 0.57$).

To determine the nonbonded parameters (partial charges and Lennard-Jones parameters) for $[H_2O]_3H^+$ 2500 reference structures were taken from an MD simulation at 300 K with reference interaction energies $E_{\text{int}} = E_{\text{tot}} - E(H_3O^+) - E(W_1) - E(W_2)$ calculated at the MP2/6-311++G(2d,2p) level of theory. Here, W_1 and W_2 are the two water molecules, respectively. All parameters from the Zundel-ion were retained and only the nonbonded parameters for the H_3O^+ ion were adjusted. The quality of this fit is shown in Figure 3b. Finally, as a validation of the parameters, MD simulations were carried out for $[H_2O]_4H^+$ at 300 K and 2500 structures were extracted. Again, interaction energies were determined at the MP2/6-311++G(2d,2p) level of theory which are then compared with MS-MMPT-MP2 energies, see Figure 3c.

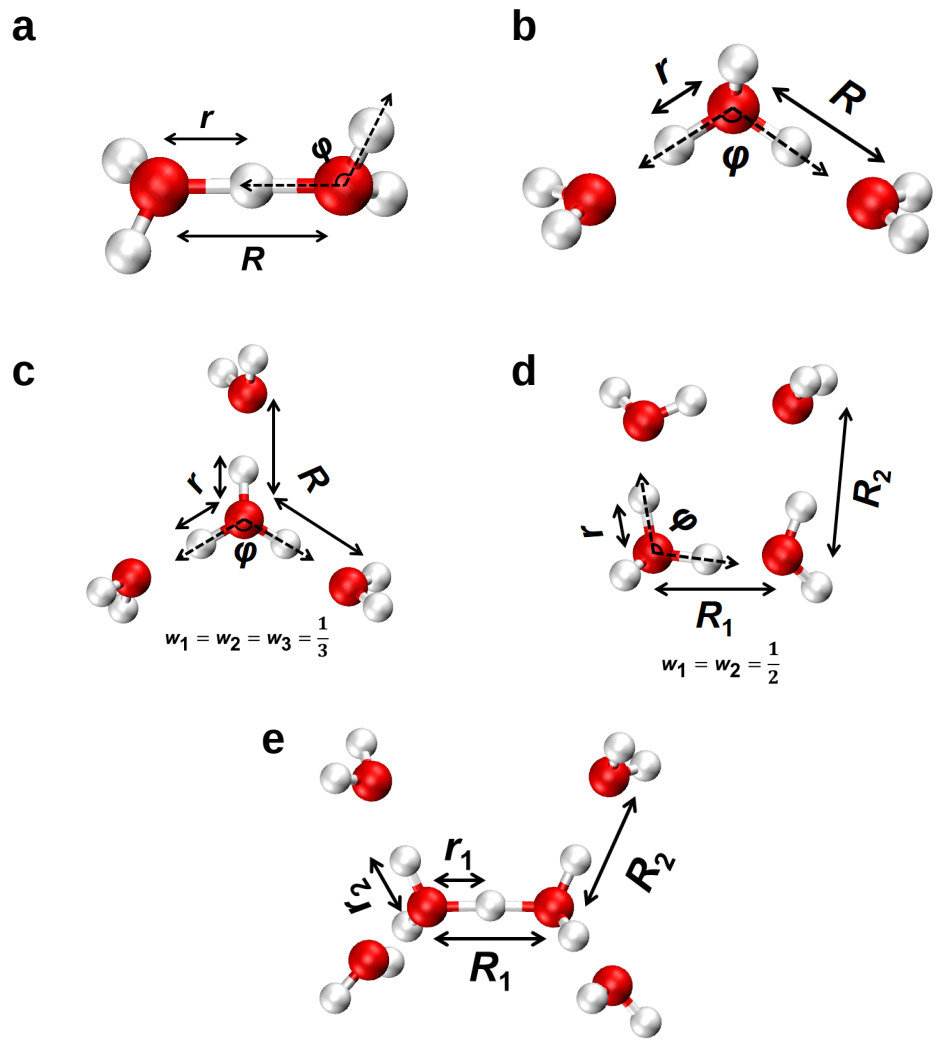


Figure 4: Geometries for optimized $[H_2O]_nH^+$ structures calculated using the MS-MMPT-MP2 force field, fitted to reference data at the MP2/6-311++G(2d,2p) level of theory. For $n = 4$ the Eigen structure consists of 3 states with weights equal to each other and the ring structure involves two states with equal weight.

3.2 Minimum energy structures of $[H_2O]_nH^+$ water clusters

Using the MS-MMPT force field, the structures of the H_2O and H_3O^+ monomers were optimized. These results agree well with the calculations at the MP2/6-311+G(2d,2p) level of theory (see Table S4), to which the MMPT force field was parametrized. Additional calculations were carried out at the CCSD(T)/aug-cc-pVTZ level of theory and the deviations were

found to be within ±1 % which establish the accuracy of the MS-MMPT-MP2 potential for the Zundel cation. For the dissociation energy, no experimental data is available. MS-MMPT-MP2 yields a value of $D_e = -33.6$ kcal/mol compared to –33.9 kcal/mol and –33.4 kcal/mol at the MP2 and CCSD(T)/aug-cc-pVTZ levels of theory, and values between –34.1 kcal/mol and –33.7 kcal/mol from complete basis set calculations at the MP2 and CCSD(T) levels of theory, respectively.[51] No weights are assigned to the H_3O^+ monomer or the $H_5O_2^+$ ion as they are parametrized explicitly.

Next, with the MS-MMPT-MP2 parametrized force field, minimum energy structures were determined for $[H_2O]_nH^+$ ($n = 2, 3, 4, 6$). Results were compared with optimized structures from MP2/6-311++G(2d,2p) calculations and are found to be in general good agreement with them. Differences for important internal coordinates (Table 1) between MS-MMPT and MP2 calculations are < 4 %. For $n = 3$ and 4, the MS-MMPT-MP2 force field yields symmetric Eigen structures in which two and three states contribute equally to the total energies, respectively. For $n = 6$, a Zundel ion is formed with energy minimization using both MP2 and MS-MMPT force field methods. That means among all candidate states only one of them is the energy minimum state which is considerably more stable than other candidate states and contributes most for describing the energetics of the system.

3.3 Improved Parametrization Based on Infrared Spectroscopy

Spectra from finite-temperature simulations: In a next step the infrared (IR) spectra for protonated water clusters were determined using MS-MMPT-MP2 from finite-temperature MD simulations (50 K with $\Delta E = 10$ kcal/mol). The IR spectrum was calculated from the Fourier transform of the dipole autocorrelation function. Vibrational spectroscopy is sensitive to the motion around the minimum of the PES and provides additional information about the curvature of the PES.

For protonated water dimer using the MMPT force field the proton transfer band is between 920 and 1090 cm^{-1} (depending on the temperature at which the simulations are run) which is consistent with accurate quantum calculations[52] on a fully dimensional potential energy surface which find this mode at 918 and 1033 cm^{-1} (compared with 928 and 1047 cm^{-1} from experiment), or 995 cm^{-1} and 1070 cm^{-1} from diffusion Monte Carlo and multimode simulations on the same 15-dimensional PES.[53] **Thus, the overall shape of the MS-MMPT-MP2 PES provides a suitable proxy to be used in further refinement.**

The $n = 4$ spectrum from MS-MMPT-MP2 simulations (Figure 5a) shows a doublet at 3800 cm^{-1} and 3880 cm^{-1} (symmetric and asymmetric water-OH stretches) compared with 3620 cm^{-1} and 3700 cm^{-1} from experiment.[54,55] The water OHO bending mode is between 1800 cm^{-1} and 1830 cm^{-1} while experimentally it is near 1600 cm^{-1}. The broad band centered at 2660 cm^{-1} from experiments is the asymmetric hydronium stretch which is located around 3440 cm^{-1} in the simulations. This shift of ~ 800 cm^{-1} is indicated by the arrow in Figure 5a and suggests that the MS-MMPT-MP2 force field can not be directly transferred from the Zundel to the Eigen ion.

To further improve the parametrization, the MS-MMPT-MP2 energy function was refined by comparing the computed with the experimentally measured IR spectrum for $[H_2O]_4H^+$: the OH stretching force constants were changed to $k_b = 250$ kcal/mol$\cdot$Å^{-2} and 550 kcal/mol$\cdot$Å^{-2} for H_3O^+ and H_2O, respectively, and the water HOH bending force constant was decreased to $k_a = 45$ kcal/mol $\cdot$ Å^{-2}. This parametrization is referred to as MS-MMPT-IR and is used in the remainder of this work.

Optimized cluster structures using MS-MMPT-IR are also reported in Table 1 and differ little from the MS-MMPT-MP2 results. This is because the two parametrizations only differ

in two force constants which do not affect equilibrium geometries appreciably. The resulting IR spectrum with the molecular dipole moment computed from the fluctuating point charges[56,57] is reported in Figure 5d.

With these refined parameters the water-OH doublet, the position of the H_3O^+ asymmetric stretch, and the water-bend are now in satisfactory agreement with experiment. Furthermore, the spectra only depend little on the choice of ΔE, see red and blue traces in Figure 5d: they slightly shift the position of the H_3O^+ asymmetric stretch but leave all other bands unaffected. For small ΔE the Zundel motif (Z) is favored, i.e. the curvature along the proton transfer coordinate is that of "pure Z" whereas for large ΔE an Eigen structure (E) is favored, see also Figure S3. Hence, the curvature along the proton transfer coordinate is a superposition of the anharmonic Zundel motif and two harmonic OH stretches with a larger curvature which yields an overall larger curvature (i.e. higher frequency) compared to pure Zundel depending on the value of ΔE. This is consistent with the notion that the anharmonicity of the proton transfer potential in MS-MMPT contributes more for larger values of ΔE.

Infrared spectra were also determined for larger clusters $(H_2O)_nH^+$ $n = 11$ and 21 (Figure 5, right panels). For $n = 11$, the PT band is well described by the simulations (2600 $\sim$ 2900 cm^{-1} compared to 2400 $\sim$ 2900 cm^{-1} from experiment). The water bending mode peaks at 1600 cm^{-1} which agrees with experiment. Similar to the spectrum for $n = 4$ from MS-MMPT-IR simulations, the water stretching bands are between 3600 $\sim$ 3800 cm^{-1}. In the experimental spectrum, however, the bonded water-OH stretches are located between 3100 cm^{-1} to 3700 cm^{-1}. Specifically, the doublet near 3700 cm^{-1} was attributed to OH stretches of "acceptor donor" (AD) and "acceptor acceptor donor" (AAD) water molecules. The sharp feature near 3600 cm^{-1} corresponds to "donor donor acceptor" (DDA) water molecules within the H-bonded network.[55] These assignments are based on power spectra of

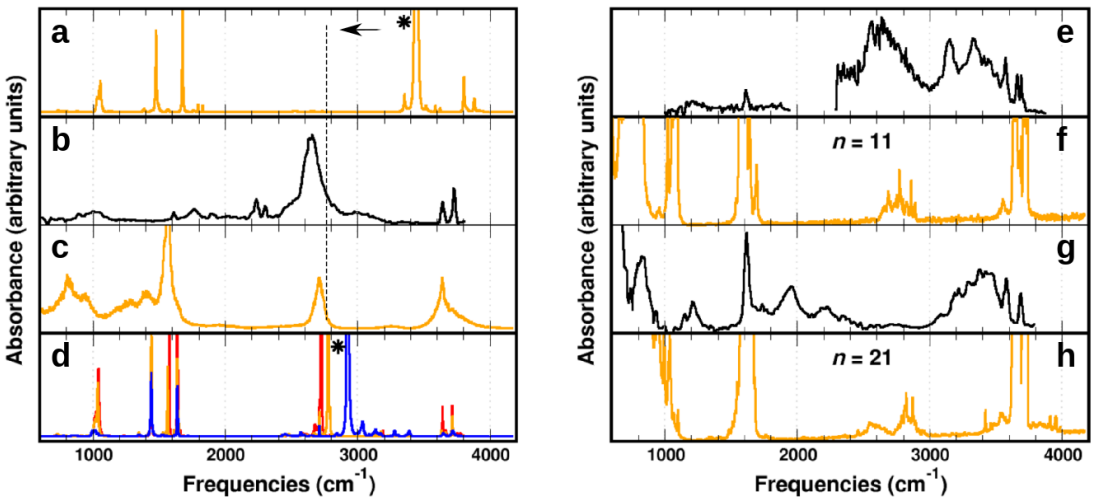


Figure 5: IR spectra of $[H_2O]_4H^+$ from a) MD simulations at 50 K with the MS-MMPT-MP2 force field (with ΔE = 10 kcal/mol), b) experimental data from Ref. 54, and MD simulations using the MS-MMPT-IR force field at c) 600 K (with ΔE = 10 kcal/mol) and d) 50 K (red: ΔE = 5 kcal/mol; orange: ΔE = 10 kcal/mol; blue: ΔE = 15 kcal/mol). The proton transfer bands from MD simulations were labeled with asterisks. The right panel shows IR spectra of $[H_2O]_{11}H^+$ from e) experiment,[55] f) MS-MMPT-IR simulations at 50 K and $[H_2O]_{21}H^+$ from g) experiment[54] and h) MS-MMPT-IR simulations at 50 K. All experimental data are from digitization.

the respective internal coordinate which has already been used in earlier work on protonated water dimer.[49] For $n = 21$, similar spectral patterns for OH stretches and water HOH bends of water molecules were found compared to those of $n = 11$. The high-frequency part of the spectrum is realistically modelled whereas the region between 3200 and 3500 cm^{-1} does not show the broad distribution observed experimentally.[54] However, recent experiments[58] and simulations[59,60] suggest that this part of the spectrum is affected by more involved couplings between overtone bending ($2\nu_b$) and OH-stretching vibrations.

There are a number of ways how the calculated IR spectra can be further improved or why the present simulations do not capture all features of the experimentally observed spectral signatures. First, all bonded interactions involving water molecules employ a harmonic en-

ergy function for the OH-stretch and the OHO-bend. Secondly, it has been recently found that some spectral features may originate from a Fermi resonance between the HOH bending overtone and the OH-stretches as in water clusters and in liquid water.[59,60] This finding in charged water clusters is quite recent;[58] it has been previously believed that these features are primarily due to OH-stretch motions.[55] Capturing a Fermi resonance is usually not possible from classical MD simulations.[38] Rather, an effective excitonic Hamiltonian is used to describe such effects.[60] Furthermore, the dipole moment function used in the present work is based on a point charge model whereas mechanical and electrical anharmonicity will affect the computed intensities. Hence, even if the vibrational frequencies are present in the underlying dynamics, the bands' intensities may be incorrectly distributed due to deficiencies in the dipole moment function. Similarly, higher multipolar moments in the charge distribution will also contribute to changes in the band positions.[61–64] Finally, realistically modeling the anharmonic coupling between the different degrees of freedom may also further improve the computed spectra.[58]

Overall, MS-MMPT-IR leads to a softer H_3O^+ ion and water molecule compared to the MS-MMPT-MP2 force field. Figure 3c (blue dots) shows the correlation between MP2 energies and those of the refined FF for $[H_2O]_4H^+$ clusters compared with the original MP2-based parametrization. The correlation coefficient decreases from $R^2 = 0.69$ to $R^2 = 0.57$. However, a direct comparison of the performance of the two parametrizations is difficult because MS-MMPT-MP2 is entirely fitted to MP2-energies (and reproduces them well) whereas MS-MMPT-IR also contains information from experiment by comparing with infrared spectra.

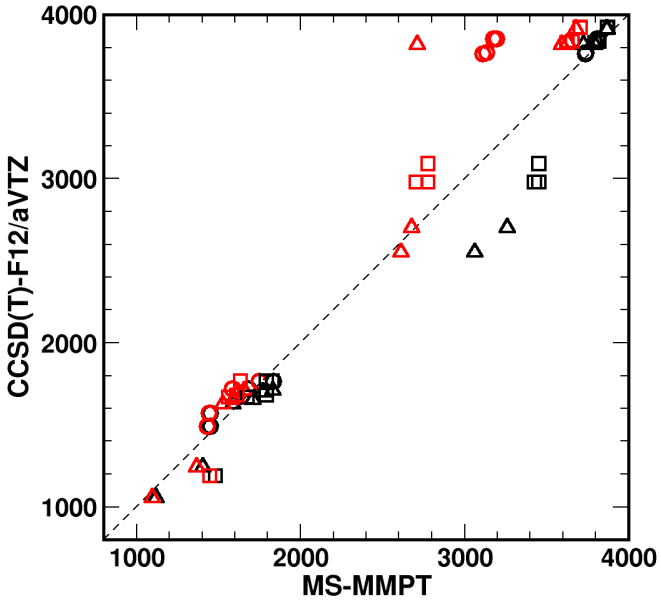


Figure 6: Comparison of harmonic frequencies for OH stretches and HOH bends between the MS-MMPT-MP2 force field (x-axis, black) and the MS-MMPT-IR force field (red), respectively, and calculations at the CCSD(T)-F12/aVTZ level of theory (y-axis)[51] at the minimum energy conformations for $[H_2O]_2H^+$ (Zundel, circle), $[H_2O]_3H^+$ (triangle), and $[H_2O]_4H^+$ (Eigen, square) structures.

Normal Modes: Although the primary aim of the present work is to introduce a computationally efficient model to follow proton dislocation in liquid water, it is instructive to compare normal mode frequencies with results from reference CCSD(T)-F12/aVTZ calculations for $[H_2O]_nH^+$ ($n = 2, 3, 4$).[51] However, no direct comparison with experiment should be made due to the harmonic approximation used. An overview of the harmonic frequencies from the two parametrizations compared with CCSD(T)-F12/aVTZ calculations is shown in Figure 6. The numerical values for the frequencies are provided in Tables S5 to S7.

Frequencies at the MP2/6-311+G(2d,2p) level of theory compare surprisingly well with results from CCSD(T)-F12/aVTZ calculations. OH-stretching frequencies are typically within

~ 10 cm^{-1} and bending frequencies differ by between 10 and 40 cm^{-1} with root mean squared differences of 3.5 cm^{-1} for the stretching and 7.8 cm^{-1} for the bending vibrations. For $H_5O_2^+$ (see Table S5) MS-MMPT-MP2 force field calculations yield results close to those from the reference MP2/6-311+G(2d,2p) calculations (and therefore to CCSD(T)-F12/aVTZ) except for one of the bending vibrations which are described by harmonic potentials.[51] Overall, results from the MS-MMPT-IR force field (red symbols in Figure 6) compare favourably with the CCSD(T)-F12/aVTZ benchmark and perform better than MS-MMPT-MP2, except for the bonded -OH stretch (red circles) of the Zundel ion which are too soft. Compared with the many-body PES, which is based on CCSD(T)/aVTZ calculations and involves a fit to ~ 8700 invariant polynomials some of which with up to 240 terms,[65] the present, more empirical model uses only ~ 30 parameters (see Tables S2 and S3) and is therefore computationally considerably more effective.

In summary, the MS-MMPT-MP2 parametrization was refined by comparing with the IR spectrum for $(H_2O)_4H^+$ to improve the force constants of a few essential modes. Comparison with experiment for larger clusters is still reasonable for $n = 11$ but fails for $n = 21$ mainly because the water OH-stretch vibration is described with a conventional harmonic potential instead of an anharmonic (Morse) term which leads to a considerably larger density of states.

3.4 Simulations for $[H_2O]_nH^+$ Clusters for $n = 6$ to 100

With the improved MS-MMPT-IR parametrization MD simulations were carried out for small to medium-sized protonated water clusters with a particular focus on the proton transfer process itself. Each simulation was 10 ns in length and was run with both force fields at temperatures of ~ 300 K using a range of ΔE values. All simulations were carried out in the NVE ensemble and the results reported below are averages over 10 independent simulations.

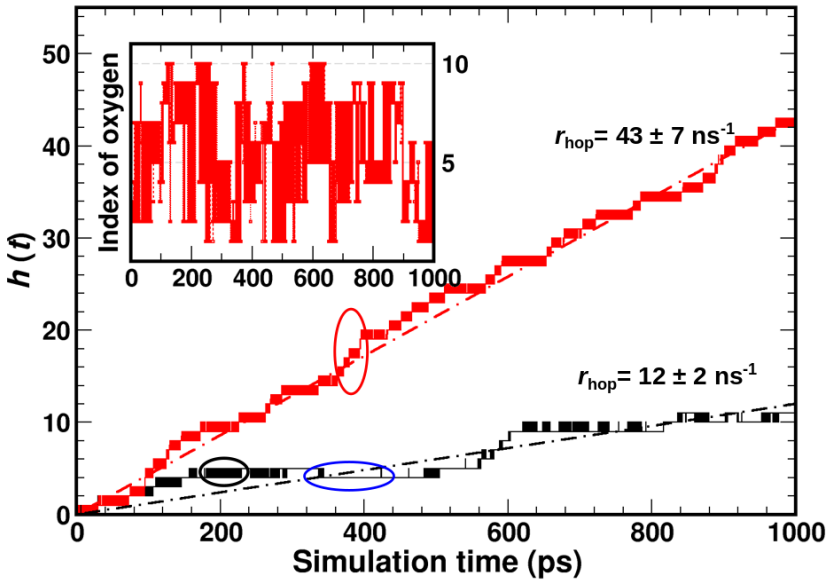


Figure 7: Time accumulation functions of proton hopping from MD simulations of water clusters with $n = 6$ (black) and 10 (red) using the MS-MMPT-IR force field with $\Delta E = 5$ kcal/mol. For each of the simulations, one representative trajectory (of which the hopping rate is closest to the trajectory average) is displayed. "Rest" and "burst" phases are found for both traces. The insets show the index profiles of the H*-carrying oxygens for a cluster simulation with $n = 10$.

As an example for the analysis of PT based on $h(t)$, Figure 7 compares the time accumulation functions for $n = 6$ and $n = 10$ using the MS-MMPT-IR force field and $\Delta E = 5$ kcal/mol. With MS-MMPT-IR and for $n = 6$ the averaged hopping rates were 12 ns^{-1} (black dot-dashed line) which increased to 43 ns^{-1} (red dot-dashed line) for $n = 10$. Previous computational studies using AIMD simulations on the 25 ps time scale had found "burst" and "resting" phases during proton transfer.[1] Using MS-MMPT-IR much more extensive sampling can be carried out (here 1 ns) which allows one to converge such time accumulation functions. A 'burst' refers to PT involving several donor atoms within a short time interval (< 10 ps), which leads to rapid increase of $h(t)$ (see red circle in Figure 7). A "resting" phase, on the other hand, corresponds to localization of the excess proton during extended

Table 2: Hopping rates for protonated water clusters from MD simulations using the MS-MMPT-IR force field for different values of ΔE.

	ΔE, kcal/mol	r_{hop}, ns^{-1}
$n = 6$	5	12 ± 2
	10	4 ± 3
	15	0
$n = 10$	5	43 ± 7
	7	69 ± 7
	10	22 ± 4
	15	1 ± 1
$n = 21$	5	56 ± 9
	7	141 ± 14
	10	45 ± 9
	15	2 ± 2
$n = 31$	5	52 ± 10
	7	130 ± 12
	10	52 ± 7
	15	3 ± 5
$n = 50$	5	55 ± 9
	7	132 ± 12
	10	49 ± 9
	15	3 ± 2
$n = 100$	5	57 ± 7
	10	131 ± 16
	12	46 ± 8
	15	3 ± 2

times. A "resting state" can be further subdivided into localization of the excess proton on the same oxygen atom (for which $h(t)$ remains unchanged, see the blue circle) and a proton repeatedly transferring within a given Zundel motif ($h(t)$ oscillating, see the black circle). The burst/rest phases can also be analyzed using index plots which follow the identity of the oxygen atom that carries the excess proton as a function of time. The H*-carrying oxygen was determined by a geometric criterion.[20] Similar to the analysis of $h(t)$ the index plots allow to identify the hopping ($\mathrm{O_A} \rightarrow \mathrm{O_B} \rightarrow \mathrm{O_C}$), resting ($\mathrm{O_A} \rightarrow \mathrm{O_A} \rightarrow \mathrm{O_A}$) and the oscillatory PT phases ($\mathrm{O_A} \rightarrow \mathrm{O_B} \rightarrow \mathrm{O_A}$). Among the three modes, only the hopping mode contributes

to the PT hopping rates.

Proton transfer rates were determined for a range of cluster sizes and ΔE values from time-course accumulation functions $h(t)$, see Table 2. It is found that for a given value of ΔE the hopping rate r_{hop} saturates for $n \geq 21$. The maximum r_{hop} depends on ΔE and ranges from 50 ns^{-1} (for $\Delta E = 5$ and 10 kcal/mol) to 130 ns^{-1} for $\Delta E = 7$ kcal/mol. Increasing ΔE to 15 kcal/mol suppresses proton transfer altogether. Similar observations are made for MS-MMPT-MP2 although the maximum rate saturates around 100 transfers/ns, see Table S8.

For the smallest clusters the maximal transfer rates are smaller compared with larger clusters (see Figure 8) because the number of available proton accepting water molecules is small which also reduces the number of candidate states. It is found that for $21 \leq n \leq 31$ the maximum rate, also representative for simulations in solution, is achieved. This suggests that "bulk-like behavior" is obtained for a finite cluster size of $\sim$25 water molecules hydrating the Zundel motif. This is consistent with earlier findings that the dynamics in the second solvation shell is essential for spatial proton transport.[16] For larger clusters the number m_{eff} of retained candidate states saturates which leads to saturation of the hopping rate. From $g_{\mathrm{H^*O_w}}(r)$ in bulk simulations the number of water oxygens to form a cluster with $n = 21$ is found to include the second maximum and corresponds to the second solvation sphere. In other words, bulk behavior is found for solvent molecules including the second hydration shell.

The dependence of the hopping rate on the value of ΔE is related to the topography of the underlying potential and free energy surfaces. Therefore, $G(q_1, q_2)$ depending on two progression coordinates was constructed for different values of ΔE from unbiased MD simulations. One progression coordinate (q_1) is the difference between the weights w_1 and w_2 of the lowest states as they distinguish between the Z- and E-states. For $|w_1 - w_2| = 1$, the state is pre-

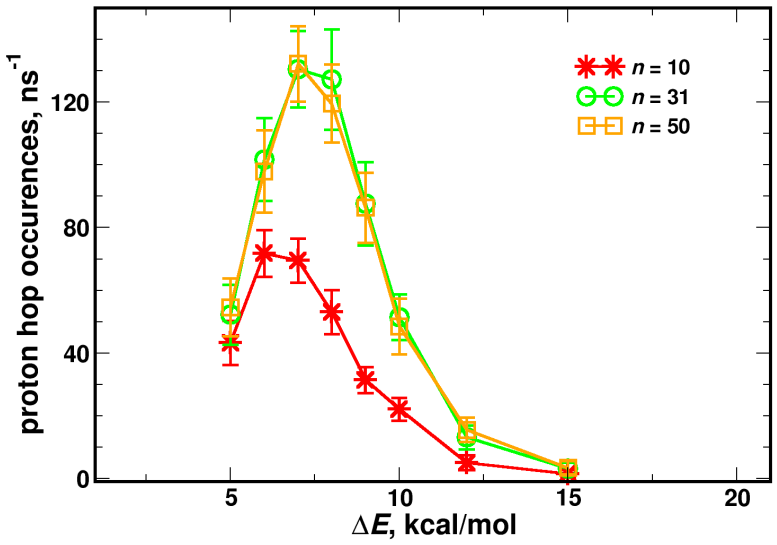


Figure 8: Dependence of the PT hopping rates on ΔE from MD simulations using the MS-MMPT-IR force field for protonated water clusters with $n = 10, 31$ and 50. Data was averaged over 10 independent trajectories, each 1 ns in length. For these simulations and $\Delta E = 7$ kcal/mol the maximum hopping rates are reached for all selected cluster sizes. For $n = 10$, the maximum rate is 69 hops/ns compared to 130 and 132 hops/ns for $n = 31$ and $n = 50$. Hence, the hopping rates depend on the cluster size but appear to converge with larger n.

dominantly that of a Zundel (Z) structure whereas for $w_1 - w_2 = 0$ (i.e. $w_1 = w_2$) there are at least two states that contribute equally, indicative of an Eigen (E) structure (see Figure 4). As a second progression coordinate q_2 the minimum δ of the displacement coordinate[20] $\delta_{\mathrm{AB}} = d(\mathrm{O_A H}*) - d(\mathrm{O_B H}*)$ and similar for δ_{AC} and δ_{AD} is considered. With these two coordinates, two-dimensional free energy surfaces (FESs) were generated from the probability distribution function, $P(w_1 - w_2, \delta)$, according to $G(w_1 - w_2, \delta) = -k_B T \log[P(w_1 - w_2, \delta)]$. For MS-MMPT-IR and $\Delta E = 5$ kcal/mol only the Z-state is a minimum on the PES whereas for $\Delta E = 15$ kcal/mol only the E-state appears, see Figure S3. For an intermediate value (such as $\Delta E = 7$ kcal/mol) both states are present as should be according to recent experiments.[7,66,67] This together with the findings for the IR spectroscopy leads us to consider

only $\Delta E \sim 7$ kcal/mol in the following.

Figure 9 shows FESs from MD simulations for $n = 31$, $n = 100$ and bulk using MS-MMPT-IR and $\Delta E = 7$ kcal/mol. The free energy surface for $n = 31$ shows a barrier height of 0.6 kcal/mol between the Z and E states which increases to 0.7 and 1.4 kcal/mol for $n = 100$ and bulk water. The similar barrier heights for the two clusters is consistent with comparable proton transfer rates of 141 ± 14 and 131 ± 16 ns^{-1}, respectively (see Table S8). The increase of the barrier for bulk water is reflected in the slowdown of the hopping rate to 60 to 70 ns^{-1}, depending on simulation conditions, see Table 3. Compared with the free energy surfaces for the two clusters that for the excess proton in bulk water exhibits an additional state, which is the "special pair",[68] discussed further below. Hence, proton diffusion and hops involve an additional state which may also contribute to the slowdown in going from the cluster to the bulk.

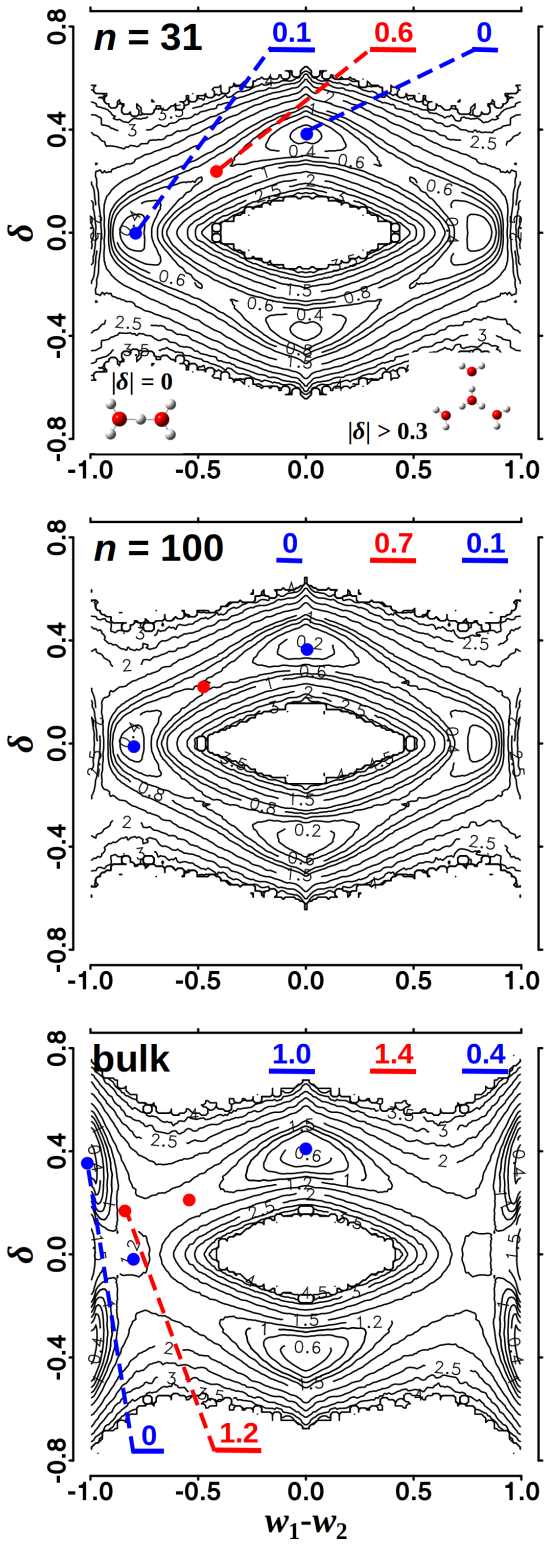


Figure 9: Two-dimensional FES for PT reactions from MD simulations with $n = 31$ and 100 and in bulk. The coordinates are δ (see the top panel for definition) and $w_1 - w_2$. A value of $|\delta| \simeq 0$ corresponds to a Zundel (Z) structure whereas $|\delta| > 0.3$ is an Eigen (E) structure.[20] MS-MMPT simulations were run for 10 ns for water clusters using the MS-MMPT-IR with $\Delta E = 7$ kcal/mol. Free energy surface of PT reactions from MS-MMPT bulk simulations using the MS-MMPT-IR force field ($\Delta E = 8$ kcal/mol, bottom). For each panel, simulations were run for 8 ns. Energies and positions of local minima (blue) and transition states (red) are indicated. For results with MS-MMPT-MP2, see Figure S4.

3.5 Proton Diffusion in Bulk Water

Next, the dynamics of the excess proton in bulk water was analyzed using MS-MMPT-IR. MD simulations using periodic boundary conditions (PBC) were carried out in the NVE and NVT ensembles. Two system sizes were considered: $25.0 \times 25.0 \times 25.0$ Å^3 (referred to as bulk-25 with 523 water molecules) and $31.0 \times 31.0 \times 31.0$ Å^3 (bulk-31 with 997 water molecules).

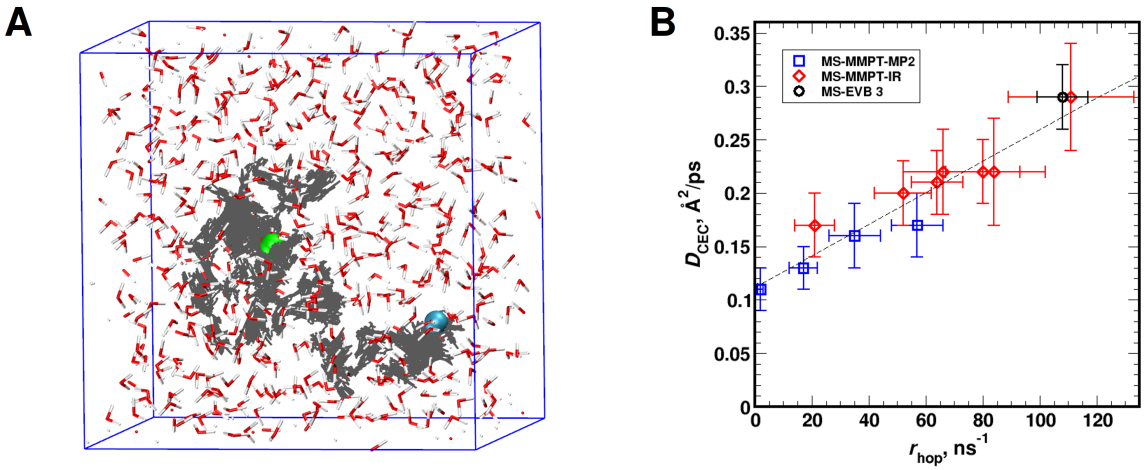


Figure 10: A) Diffusion trajectory (250 ps) of the center of excess charge in bulk-25 simulation and B) Relationship between hopping rate r and D_{CEC} including error bars. In the panel A, the diffusion trace is colored in gray and the green and blue dots are the starting and end point, respectively. In the panel B, the result from MS-EVB3[29] is indicated in black.

The self-diffusion coefficient of the excess proton was determined from following the center of excess charge (CEC)[28,29]

$$\vec{\boldsymbol{r}}_{\mathrm{CEC}} = \sum_{j=1}^{N} w_j \cdot \vec{\boldsymbol{r}_j} \tag{14}$$

where w_j is the weight of the j-th state and $\vec{\boldsymbol{r}_j}$ is the coordinate of weighted center of excess charge

$$\vec{\boldsymbol{r}_j} = (1-\gamma) \cdot \frac{\sum\limits_k |q_k^{\mathrm{I}}| \cdot r_k^{\mathrm{I}}}{\sum\limits_k |q_k^{\mathrm{I}}|} + \gamma \cdot \frac{\sum\limits_k |q_k^{\mathrm{II}}| \cdot r_k^{\mathrm{II}}}{\sum\limits_k |q_k^{\mathrm{II}}|} \tag{15}$$

with $q_k^{\mathrm{I/II}}$ and $r_k^{\mathrm{I/II}}$ the charges and positions of atoms in the hydronium ion described in the two resonance structures. Figure 10 shows one trajectory following the CEC from a 250 ps simulation using MS-MMPT-IR which shows that extensive proton diffusion occurs. From $\vec{\boldsymbol{r}}_{\mathrm{CEC}}$ the self-diffusion coefficient is obtained from its mean square displacement

$$D(t) =< |\vec{\boldsymbol{r}}_{\mathrm{CEC}}(t) - \vec{\boldsymbol{r}}_{\mathrm{CEC}}(0)|^2 > / \, 6t. \tag{16}$$

It is important to note here that in MS-MMPT no individual hydronium ion serves as a proton carrier. Rather, a Zundel motif as a whole carries the proton.

Hopping rates are again determined from time accumulation functions. Table 3 shows that the hopping rates in bulk simulations also depend on the value of ΔE, as was the case for the clusters (see Table 2). They range from 20 ns^{-1} to 111 ns^{-1} depending on box size and nonbonded cutoff. For the bulk-31 simulations where longer nonbonded cutoffs can be used the hopping rates and diffusion coefficients D_{CEC} increase by 40 % to 50 % compared to simulations in the 25 Å box. The values for r_{hop} and D_{CEC} for the MS-MMPT-IR simulations are in general larger than those for simulations with MS-MMPT-MP2 (see Table S9) and are close to those for the $n = 100$ cluster simulations (i.e. 130 hops/ns). The diffusivity of the center of excess charge approaches a value of 0.30 $\mathrm{Å}^2 \cdot \mathrm{ps}^{-1}$. Figure 10 reports the relationship between D_{CEC} and the hopping rate. It can be seen that with increasing hop rate the diffusivity increases, too, but at a slower rate.

In a typical simulation, the average number of candidate states is 16.7 for first-shell motifs and increases to 110.0 if water molecules in the second shell are considered in the state selection (see Figure S5). Application of the weight criterion ($w^i \leq 10^{-11}$) to these states to retain them for computing the total potential energy of a configuration leaves an average number of 8.3 and 13.0 when including first- and second-shell solvent molecules, respectively. Including

second-shell motifs is essential for conservation of total energy (see Figure S6) in particular when the identity of the minimum energy state changes. Increasing the time step to $\Delta t = 0.5$ fs leads to an energy drift of 0.0025 (kcal/mol)/ns/molecule, compared with 0.0025 (kcal/mol)/ns/molecule from MS-EVB 3.2 simulations and 0.0028 (kcal/mol(/ns/molecule for aMS-EVB 3.2.[69]

Table 3: Self-diffusion coefficients ($Å^2 \cdot ps^{-1}$) from bulk simulations with two box sizes using the MS-MMPT-IR force field.

System size	ΔE kcal/mol	Δt fs	R_{cutoff} Å	ensemble	r_{hop} ns^{-1}	D_{CEC} $Å^2 \cdot ps^{-1}$
25.0 Å	5	0.5	10	*NVE*	21(7)	0.17(3)
	7	0.5	10		64(9)	0.21(3)
	8	0.5	10		80(13)	0.22(3)
	10	0.5	10		52(10)	0.20(3)
	7	0.5	10	*NVT*	66(14)	0.22(4)
	8	0.5	10		84(18)	0.22(5)
31.0 Å	8	0.5	14	*NVE*	111(22)	0.29(5)

Similar to the cluster simulations, Figure 9 (bottom row) reports $G(q_1, q_2)$ for bulk simulations using MS-MMPT-IR. The FES exhibits three minima - the expected Z and E states and an additional third state for which the geometry (reflected in the value of δ) is that of E but the weights of the two lowest states are those of a Z state. This is the "special pair" (for a geometrical analysis of the special pair see Figure S8) known from previous MS-EVB simulations[18,68,70] which is slightly lower in energy (by 0.4 kcal/mol) than the E-structure and considerably stabilized (by 1 kcal/mol) relative to Z. The special pair is separated by a free energy barrier of 1.2 and 1.4 kcal/mol from the other two minima, respectively. The MS-MMPT-MP2 free energy surface (Figure S4) does not show this special pair. Using MS-EVB3 the E-state was reported to be more stable than the Z-state by ~ 0.6 kcal/mol (2.4 kJ/mol),[18] close to the present results. Similarly, with MS-EVB3.2 the E-form was

found to be more stable than the Z-state, by between ~ 0.5 and ~ 1 kcal/mol, depending on the progression coordinate used.[69] The relative stabilization of and barrier between the "special pair" and the Z-state have also been determined from free energy simulations using the DFTB3-diag+gaus semiempirical method. With δ as the progression coordinate, it was found that the Z-state is destabilized by ~ 0.4 kcal/mol and separated from the "special pair" by a barrier of ~ 0.9 kcal/mol. These results are in line with the present findings, although the Z-state from MS-MMPT-IR is somewhat more destabilized. Similarly, CPMD-HCTH simulations reported a stabilization of the "special pair" versus the Z-state, however, without a barrier between the two.[70,71]

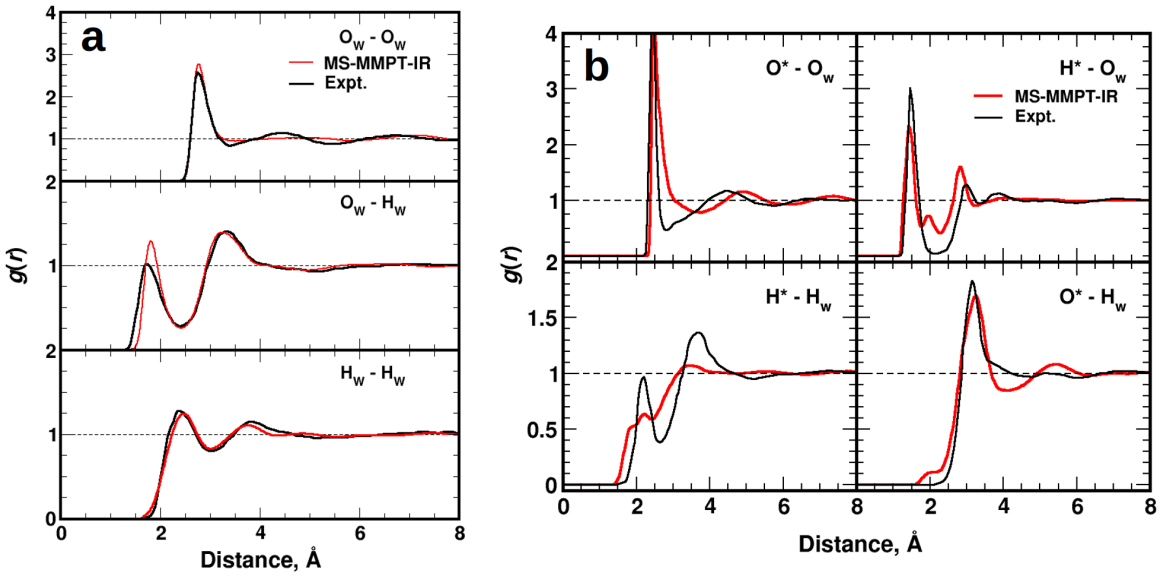


Figure 11: Radial distribution functions of a) bulk water for (from top to bottom) $g_{\mathrm{O_wO_w}}(r)$, $g_{\mathrm{O_wH_w}}(r)$ and $g_{\mathrm{H_wH_w}}(r)$ and b) protonated bulk water, including $g_{\mathrm{O^*O_w}}(r)$, $g_{\mathrm{O^*H_w}}(r)$, $g_{\mathrm{H^*H_w}}(r)$ and $g_{\mathrm{H^*O_w}}(r)$ from simulations with MS-MMPT-IR ($\Delta E = 8$ kcal/mol), compared with experiment.[72] $\mathrm{O^*}$ is the proton-carrying oxygen and $\mathrm{H^*}$ is the excess proton. For bulk water (panel a) the radial distribution functions from MS-MMPT simulations are almost indistinguishable and capture some of the experimentally measured features,[73] notably the position of the most pronounced maxima except for the first maximum of $g_{\mathrm{O_wH_w}}(r)$. The second maximum for $g_{\mathrm{O_wO_w}}(r)$ is too faint and occurs at too large a distance r.

As a final experimental observable, the water-water and water-hydronium radial distribu-

tion functions were determined (see Figure 11). First, the $g_{O_wO_w}(r)$ (water–water oxygen), $g_{O_wH_w}(r)$ (water oxygen–hydrogen) and $g_{H_wH_w}(r)$ (water–water hydrogen) radial distribution functions were obtained from 8 ns simulations in which the Zundel ion was removed. The computed $g(r)$ are compared with data from combined X-ray/neutron scattering experiments.[73] For $g_{O_wO_w}(r)$ the first density peak appears at $r = 2.75$ Å which agrees well with experiment. The experimental data shows a second peak at 4.5 Å, which is shifted to longer separations and less pronounced. For $g_{O_wH_w}(r)$ the first maximum occurs at $r = 1.81$ Å from MD simulations, compared to 1.74 Å from the experiment and the second peak at 3.28 Å is well reproduced from both force fields. Similar agreement is found for $g_{H_wH_w}(r)$.

The RDFs were also calculated from simulations of the hydrated Zundel ion. From the 8 ns simulations the $g_{O^*O_w}(r)$ (hydronium–water oxygen), $g_{O^*H_w}(r)$ (hydronium oxygen–water hydrogen), $g_{H^*H_w}(r)$ (hydronium–water hydrogen) and $g_{H^*O_w}(r)$ (hydronium hydrogen–water oxygen) were determined and compared with experiment.[72] In all cases the position of the first peak agrees well with experiment. The small feature around 2 Å in $g_{O^*H_w}(r)$, not present in the experiments, has also been found in recent simulations using MS-EVB3.2[69] **and was related to the presence of weakly bound water molecules near the hydronium ion, which are also present here, see Figure S7. It should, however, be noted, that the MS-EVB 3.2 potential was modified to reproduce this feature at ~ 2.0 Å in $g_{O^*H_w}(r)$, whereas for MS-MMPT-IR this peak is present without explicitly accounting for.** The present RDFs can also be compared with those from DFTB-diag+gaus simulations. The $g_{H^*O_w}(r)$ distribution function from DFTB-diag+gaus has a double-maximum at short range which differs from experiment, MS-MMPT-IR, and MS-EVB whereas between 2.0 and 2.5 Å the RDFs from DFTB-diag+gaus and MS-MMPT-IR are non-zero which differs from experiment and MS-EVB. For $r > 3.0$ Å this RDF from all methods are comparable. The height and position of the first maximum of $g_{O^*H_w}(r)$ compared with experiment is well captured by MS-MMPT-IR and MS-EVB whereas DFTB-

diag+gaus overestimates the amplitude and the maximum is shifted to somewhat larger r. The first minimum with MS-MMPT-IR and DFTB-diag+gaus is at $r \approx 4.0$ Å whereas that with MS-EVB is at $r \approx 3.5$ Å but this $g(r)$ from experiment is unstructured in this region.

The difference between the calculated and the experimental $g_{\mathrm{H}^*\mathrm{O_w}}(r) \sim 2$ Å was found to also depend on the van der Waals radius of the O* (proton carrying oxygen) atom. Reducing this from 1.74 Å to 1.65 Å the maximum in $g_{\mathrm{H}^*\mathrm{O_w}}(r)$ disappears and the shapes for $g_{\mathrm{O}^*\mathrm{O_w}}(r)$ and $g_{\mathrm{H}^*\mathrm{H_w}}(r)$ considerably changes, as demonstrated by Figure S9. Furthermore, the shape of $g_{\mathrm{O}^*\mathrm{O_w}}(r)$ agrees better with experiment with smaller van der Waals radius of O*.

For comparing experiments and the present simulations it should be noted that in the experiments[72] the proton concentration is 50 times larger (1:10 for proton:water) than that used in the present simulations ($\sim 1:500$). Furthermore, experiments were carried out for concentrated HCl solutions, i.e. the counterions were present, and the experimental results (composite partial structure factors) need to be first processed by Monte Carlo simulations using an empirical potential structure refinement.[72] Hence, the differences between the computed and experimentally observed $g(r)$ are not too surprising and it would, in fact be problematic if the $g(r)$ from experiment and simulations did agree given the differences in system and data processing between the two approaches.

4 Discussion and Conclusions

The present work introduces a computational model to study proton diffusion in bulk water based on a bottom-up approach starting with extensive reference data from MP2/6-311++G(2d,2p) calculations (MS-MMPT-MP2), as is also done in MS-ARMD[34,35,41] or MS-RMD.[74] Further refinement is based on comparison with experimental infrared experiments

(MS-MMPT-IR). Proton translocation is possible through a combination of an explicitly parametrized MMPT force field (for PT between two neighboring water molecules in a Zundel ion) combined with multi state reactive MD. Water molecules within the first two solvation shells of the proton-carrying motif need to be included for satisfactory energy conservation.

The cluster simulations also provide an opportunity to consider the question whether the excess proton is rather at an interior location or prefers a surface position. Results from a typical 1 ns simulation for $(H_2O)_{31}H^+$ are reported in Figure S10 and demonstrate that the distance of the charge defect from the center of mass of the cluster corresponds typically to the radius of gyration (average size) of the cluster. Hence, the defect prefers a location close to or on the surface. Whenever inside positions are sampled, they can not be stabilized (situation at $t = 400$ ps). This agrees with previous findings using an EVB model[75] which reported that surface positions are favoured by ~ 4 $k_B T$ compared to inside positions.

The present approach can also be compared with multi state EVB.[26,69] For the proton hopping rates the present simulations (100 to 130 ns^{-1}) support results from MS-EVB3 simulations for which 108 hops/ns (see Figure 10) were found[29] whereas semiempirical DFTB, which underestimates the barrier for proton transfer, and CPMD simulations find between 400 and 500 hops/ns.[29] Given the entirely independent simulation technology and parametrization scheme the consistency of MS-MMPT-IR and MS-EVB3 lends support for a slower proton exchange rate (~ 100 ns^{-1}) than that found from ab initio (semiempirical or DFT) simulations. Based on the good agreement with available experimental data and existing results from MS-EVB3 we consider MS-MMPT-IR with $\Delta E = 7$ kcal/mol to be the best model so far.

There are several possible improvements to the present parametrization. One concerns the parametrization of the solvent water molecules. Instead of harmonic OH-bonds an anhar-

monic parametrization akin to a Kumagai, Kawamura and Yokokawa potential[76,77] could be considered for improved spectroscopic properties. Furthermore, recent work on the terahertz and Raman spectroscopy of pure water suggests that small amounts of charge transfer between H-bonded water molecules further improve the spectroscopy and dynamics of water.[57] Finally, due to the strong local electric fields around a Zundel ion, polarization effects are expected to be appreciable and including them should further improve the model.[7]

The dynamics on the present MS-MMPT parametrizations was followed by using classical MD simulations. However, quantum effects have also been included for quite some time.[20,78,79] Using centroid MD simulations and the MS-EVB3.2 parametrization, one of the findings is that the self diffusion coefficient increases from ~ 0.3 Å^2/ps to ~ 0.5 Å^2/ps, which is an increase of about 70 %. This general finding is also consistent with recent ab initio ring polymer MD simulations including quantum effects, which report that including quantum effects generally increases the diffusivity of the charge defect.[80] However, this work also found that the actual diffusion coefficient depends sensitively on the density functional chosen and it was concluded that ".. it is difficult to gather enough statistics to estimate the diffusion coefficient from ab initio simulations using density functional theory, and when this is done the results turn out to be very sensitive to the choice of density functional."[80] Depending on the system considered, the diffusion coefficients ranged from 0.16 to 0.49 Å^2/ps. On the other hand, the influence of quantum effects on the radial distribution functions was found to be rather small.[81] Hence, depending on the property considered, including quantum effects may be required for semi-quantitative agreement with experiment.

With respect to simulation technology, EVB[23] requires diagonalization of a matrix with valence bond states as diagonal elements and off-diagonal elements that couple them at every time step. Conversely, MS-ARMD uses a linear combination of a number of candidate states. The off-diagonal elements in EVB have led to some debate[82] as they need to be parametrized

in one way or another which often involves their dependence on geometrical coordinates. In MS-MMPT the energy mixing parameter ΔE is the only free parameter once the force field for the gas phase and hydrated Zundel ion are parametrized. From a conceptual viewpoint MS-MMPT mixes the energies of candidate PT motifs (see Eqs. 10 and 11). This differs from MS-EVB which is rather based on the concept of charge delocalization[69] which, in MS-MMPT, is a consequence rather than part of the model.

In terms of computational efficiency, MS-EVB requires one to solve an eigenvalue problem at every MD step and all states and coupling elements need to be computed. Contrary to that, in MS-MMPT a relatively large number of candidate states (~ 100 in the present case) are considered which reduce to 10 to 20 on average for which the full force field energy needs to be determined. With regard to parametrization, MS-MMPT follows a bottom-up strategy whereas MS-EVB is rather more a top-down approach.

Given the considerable differences between MS-MMPT and MS-EVB3.2 it is gratifying to note that the two approaches arrive at similar conclusions, specifically concerning the hopping rates in bulk and the radial distribution functions although, specifically for $g_{\mathrm{H^*H_w}}(r)$, MS-EVB 3.2 compares better with experiment than MS-MMPT-IR. However, as already noted above, comparison with experiments needs to be done with circumspection as they use a higher proton concentration and counterion effects may lead to additional differences. Further refinements (see Figure S9) of MS-MMPT-IR by optimizing the nonbonded parameters are possible but outside the scope of the present work. Also, the important role of the "special pair" is found with both approaches which lends credibility as to its relevance for a molecular-level description of the excess proton in water.

Recently, machine learning has emerged as a potentially interesting approach to represent high-dimensional (reactive) potential energy surfaces.[83,84] One of the

hallmarks of such approaches is that they are very data-intensive. For a recent reactive MD study for the Diels-Alder reaction involving 3-dibromo-1,3-butadiene and maleic anhydride a NN based on ~ 225000 structures was trained.[85] The evaluation time for this NN was a factor of ~ 200 slower compared with a representation based on MS-ARMD which is the technique also used here. For a condensed phase system, such as the excess proton in bulk water where water molecules up to the second solvation shell $(H_2O)_nH^+$ with $n \sim 20$) need to be included, the number of the electronic structure calculations will be considerably larger due to the flexibility and structural heterogeneity of the system. Hence, despite the potential benefits of such data-driven approaches, their use for reactions in solution, which involves even greater conformational variability than the Diels-Alder reaction, is still not routine.

Another advantage of MS-MMPT is the fact that it can be easily extended to chemically heterogeneous systems as the method is entirely based on site energies, i.e. the total energy of the system depends only on where the excess proton is positioned. As an example, for the protonation/deprotonation of acetic acid in water, in addition to the symmetric double minimum (SDM) PES only the asymmetric double minimum (ADM) PES together with limited reparametrization of the equilibrium positions and barrier heights needs to be included in the simulation. This is very different for an NN-trained force field for which every change in the chemistry of the system requires learning another NN based on a new, potentially extensive, set of reference electronic structure calculations. However, in the future, techniques such as transfer learning[86] may provide computationally advantageous routes to addressing this problem.

The analysis of the free energy surfaces underlines early findings which reported "...the com-

plex behavior of the hydrated proton, bringing out features of both Eigen's and Zundel's views, but it also emphasizes that the defect cannot be fully understood using either of these...".[20] More recent spectroscopic work[7,66,67] reported that Zundel-type hydration geometries are prevalent in bulk water. This is consistent with the present findings using the MS-MMPT-IR parametrization which locate the "special pair", corresponding to a hydrated Zundel structure, as the minimum on the free energy surface, separated by 1.2 kcal/mol from the Zundel form itself. **This is in good agreement with recent experiments which report a barrier for proton transfer of ~ 1 kcal/mol.[87]** Because the barrier between the two Zundel and the Eigen form is higher (1.4 kcal/mol) it is expected that the dynamics of the excess proton in bulk water is dominated by Zundel-like forms. This is also supported by recent experiments which found that the spectroscopic signatures for proton transfer are strongly IR active and weakly Raman active, indicative of an approximately centrosymmetric Zundel-like structure as the main spectroscopically responsive species.[67]

Supporting Information

The supporting information contains detailed illustration to the MS-MMPT method and the MMPT potential for single proton transfer. Parameters of the MS-MMPT force field, normal mode analysis for small hydrated proton clusters and more details of MS-MMPT simulation for clusters and water bulks with one excess proton are also provided. All materials are available free of charge via the Internet at http://pubs.acs.org/.

Acknowledgments

The authors gratefully acknowledge financial support from the Swiss National Science Foundation through the NCCR MUST and grant 200021-117810.

References

(1) Hassanali, A.; Giberti, F.; Cuny, J.; Kühne, T. D.; Parrinello, M. Proton Transfer through the Water Gossamer. *Proc. Natl. Acad. Sci.* **2013**, *110*, 13723–13728.

(2) Marx, D. Proton Transfer 200 Years after Von Grotthuss: Insights from Ab Initio Simulations. *Comput. Phys. Commun.* **2007**, *8*, 209–210.

(3) Marx, D.; Chandra, A.; Tuckerman, M. E. Aqueous Basic Solutions: Hydroxide Solvation, Structural Diffusion, and Comparison to the Hydrated Proton. *Chem. Rev.* **2010**, *110*, 2174–2216.

(4) Masgrau, L.; Roujeinikova, A.; Johannissen, L. O.; Hothi, P.; Basran, J.; Ranaghan, K. E.; Mulholland, A. J.; Sutcliffe, M. J.; Scrutton, N. S.; Leys, D. Atomic Description of an Enzyme Reaction Dominated by Proton Tunneling. *Science* **2006**, *312*, 237–241.

(5) Nagle, J. F.; Morowitz, H. J. Molecular Mechanisms for Proton Transport in Membranes. *Proc. Natl. Acad. Sci.* **1978**, *75*, 298–302.

(6) Morsomme, P.; Boutry, M. The Plant Plasma Membrane H^+-ATPase: Structure, Function and Regulation. *Biochim. Biophys. Acta* **2000**, *1465*, 1–16.

(7) Dahms, F.; Fingerhut, B. P.; Nibbering, E. T. J.; Pines, E.; Elsaesser, T. Large-Amplitude Transfer Motion of Hydrated Excess Protons Mapped by Ultrafast 2D IR Spectroscopy. *Science* **2017**, *357*, 491–494.

(8) Giguere, P. A.; Turrell, S. The Nature of Hydrofluoric Acid. A Spectroscopic Study of the Proton-Transfer Complex $H_3O^+F^-$. *J. Am. Chem. Soc.* **1980**, *102*, 5473–5477.

(9) Kim, J.; Schmitt, U. W.; Gruetzmacher, J. A.; Voth, G. A.; Scherer, N. E. The Vibrational Spectrum of the Hydrated Proton: Comparison of Experiment, Simulation, and Normal Mode Analysis. *J. Chem. Phys.* **2002**, *116*, 737–746.

(10) Wolke, C. T.; Fournier, J. A.; Dzugan, L. C.; Fagiani, M. R.; Odbadrakh, T. T.; Knorke, H.; Jordan, K. D.; McCoy, A. B.; Asmis, K. R.; Johnson, M. A. Spectroscopic Snapshots of the Proton-Transfer Mechanism in Water. *Science* **2016**, *354*, 1131–1135.

(11) Fournier, J. A.; Johnson, C. J.; Wolke, C. T.; Weddle, G. H.; Wolk, A. B.; Johnson, M. A. Vibrational Spectral Signature of the Proton Defect in the Three-Dimensional $H^+(H_2O)_{21}$ Cluster. *Science* **2014**, *344*, 1009–1012.

(12) Mackeprang, K.; Xu, Z.-H.; Maroun, Z.; Meuwly, M.; Kjaergaard, H. G. Spectroscopy and Dynamics of Double Proton Transfer in Formic Acid Dimer. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24654–24662.

(13) Wolke, C. T.; DeBlase, A. F.; Leavitt, C. M.; McCoy, A. B.; Johnson, M. A. Diffuse Vibrational Signature of a Single Proton Embedded in the Oxalate Scaffold, $HO_2CCO_2^-$. *J. Phys. Chem. A* **2015**, *119*, 13018–13024.

(14) Howard, D. L.; Kjaergaard, H. G.; Huang, J.; Meuwly, M. Infrared and Near-Infrared Spectroscopy of Acetylacetone and Hexafluoroacetylacetone. *J. Phys. Chem. A* **2015**, *119*, 7980–7990.

(15) Choi, P.; Jalani, N. H.; Datta, R. Thermodynamics and Proton Transport in Nafion: II. Proton Diffusion Mechanisms and Conductivity. *J. Electrochem. Soc.* **2005**, *152*, 123–130.

(16) Ando, K.; Hynes, J. Molecular Mechanism of HCl Acid Ionization in Water: Ab Initio Potential Energy Surfaces and Monte Carlo Simulations. *J. Phys. Chem. B* **1997**, *101*, 10464–10478.

(17) Lapid, H.; Agmon, N.; Petersen, M.; Voth, G. A Bond-Order Analysis of the Mechanism for Hydrated Proton Mobility in Liquid Water. *J. Chem. Phys.* **2005**, *122*, 014506.

(18) Markovitch, O.; Chen, H.; Izvekov, S.; Paesani, F.; Voth, G. A.; Agmon, N. Special Pair Dance and Partner Selection: Elementary Steps in Proton Transport in Liquid Water. *J. Phys. Chem. B* **2008**, *112*, 9456–9466.

(19) Carpenter, W. B.; Fournier, J. A.; Lewis, N. H. C.; Tokmakoff, A. Picosecond Proton Transfer Kinetics in Water Revealed with Ultrafast IR Spectroscopy. *J. Phys. Chem. B* **2018**, *122*, 2792–2802.

(20) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. The Nature of the Hydrated Excess Proton in Water. *Nature* **1999**, *397*, 601–604.

(21) Luduea, G. A.; Kühne, T. D.; Sebastiani, D. Mixed Grotthuss and Vehicle Transport Mechanism in Proton Conducting Polymers from Ab initio Molecular Dynamics Simulations. *Chem. Mater.* **2011**, *23*, 1424–1429.

(22) Chen, M.; Zheng, L.; Santra, B.; Ko, H.-Y.; DiStasio, R. A., Jr.; Klein, M. L.; Car, R.; Wu, X. Hydroxide Diffuses Slower than Hydronium in Water Because Its Solvated Structure Inhibits Correlated Proton Transfer. *Nat. Chem.* **2018**, *10*, 413–419.

(23) Warshel, A.; Weiss, R. An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. *J. Am. Chem. Soc.* **1980**, *102*, 6218–6226.

(24) Warshel, A.; Weiss, R. Empirical Valence Bond Calculations of Enzyme Catalysis. *Ann. N. Y. Acad. Sci.* **1981**, *367*, 370–382.

(25) Walbran, S.; Kornyshev, A. A. Proton Transport in Polarizable Water. *J. Chem. Phys.* **2001**, *114*, 10039–10048.

(26) Schmitt, U.; Voth, G. Multistate Empirical Valence Bond Model for Proton Transport in Water. *J. Phys. Chem. B* **1998**, *102*, 5547–5551.

(27) Lefohn, A. E.; Ovchinnikov, M.; Voth, G. A. A Multistate Empirical Valence Bond

Approach to a Polarizable and Flexible Water Model. *J. Phys. Chem. B* **2001**, *105*, 6628–6637.

(28) Day, T.; Soudackov, A. V.; Cuma, M.; Schmitt, U. W.; Voth, G. A. A Second Generation Multistate Empirical Valence Bond Model for Proton Transport in Aqueous Systems. *J. Chem. Phys.* **2002**, *117*, 5839–5849.

(29) Wu, Y.; Chen, H.; Wang, F.; Paesani, F.; Voth, G. A. An Improved Multistate Empirical Valence Bond Model for Aqueous Proton Solvation and Transport. *J. Phys. Chem. B* **2008**, *112*, 467–482.

(30) Chen, C.; Arntsen, C.; Voth, G. A. Development of Reactive Force Fields using ab initio Molecular Dynamics Simulation Minimally biased to Experimental Data. *J. Chem. Phys.* **2017**, *147*, 161719.

(31) Arntsen, C.; Chen, C.; Voth, G. A. Reactive Molecular Dynamics Models From Ab Initio Molecular Dynamics Data Using Relative Entropy Minimization. *Chem. Phys. Lett.* **2017**, *683*, 573–578.

(32) Glowacki, D. R.; Orr-Ewing, A. J.; Harvey, J. N. Non-equilibrium Reaction and Relaxation Dynamics in a Strongly Interacting Explicit Solvent: F + CD3CN Treated with a Parallel Multi-State EVB Model. *J. Chem. Phys.* **2015**, *143*, 044120.

(33) Lammers, S.; Lutz, S.; Meuwly, M. Reactive Force Fields for Proton Transfer Dynamics. *J. Comput. Chem* **2008**, *29*, 1048–1063.

(34) Danielsson, J.; Meuwly, M. Atomistic Simulation of Adiabatic Reactive Processes Based on Multi-State Potential Energy Surfaces. *J. Chem. Theo. Comp.* **2008**, *4*, 1083–1093.

(35) Nagy, T.; Yosa Reyes, J.; Meuwly, M. Multisurface Adiabatic Reactive Molecular Dynamics. *J. Chem. Theo. Comp.* **2014**, *10*, 1366–1375.

(36) Meuwly, M. Reactive Molecular Dynamics: From Small Molecules to Proteins. *WIREs* **2019**, *9*, e1386.

(37) Yang, Y.; Meuwly, M. A Generalized Reactive Force Field for Nonlinear Hydrogen Bonds: Hydrogen Dynamics and Transfer in Malonaldehyde. *J. Chem. Phys.* **2010**, *133*, 064503.

(38) Xu, Z.-H.; Meuwly, M. Vibrational Spectroscopy and Proton Transfer Dynamics in Protonated Oxalate. *J. Phys. Chem. A* **2017**, *121*, 5389–5398.

(39) Brooks, B.; Bruccoleri, R.; Olafson, D.; States, D.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comp. Chem.* **1983**, *4*, 187–217.

(40) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. et al. CHARMM: The Biomolecular Simulation Program. *J. Comp. Chem.* **2009**, *30*, 1545–1614.

(41) Nutt, D. R.; Meuwly, M. Studying Reactive Processes with Classical Dynamics: Rebinding Dynamics in MbNO. *Biophys. J.* **2006**, *90*, 1191–1201.

(42) Denis-Alpizar, O.; Bemish, R. J.; Meuwly, M. Reactive Collisions for NO((II)-I-2) + N(S-4) at Temperatures Relevant to the Hypersonic Flight Regime. *Phys. Chem. Chem. Phys.* **2017**, *19*, 2392–2401.

(43) Rocha, C. M. R.; Varandas, A. J. C. Energy-Switching Potential Energy Surface for Ground-State C_3. *Chem. Phys. Lett.* **2018**, *700*, 36–43.

(44) Yosa Reyes, J.; Brickel, S.; Unke, O. T.; Nagy, T.; Meuwly, M. HSO3Cl: a prototype molecule for studying OH-stretching overtone induced photodissociation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 6780–6788.

(45) Brickel, S.; Meuwly, M. OH-Stretching Overtone Induced Dynamics in HSO3F from Reactive Molecular Dynamics Simulations. *J. Phys. Chem. A* **2017**, *121*, 5079–5087.

(46) Soloviov, M.; Das, A. K.; Meuwly, M. Structural Interpretation of Metastable States in Myoglobin-NO . *Ang. Chem. Intern.* **2016**, *55*, 10126–10130.

(47) Wu, Y.; Tepper, H. L.; Voth, G. A. Flexible Simple Point-Charge Water Model with Improved Liquid-State Properties. *J. Chem. Phys.* **2006**, *124*, 024503.

(48) Meuwly, M.; Hutson, J. Morphing ab Initio Potentials: A Systematic Study of NeHF. *J. Chem. Phys.* **1999**, *110*, 8338–8347.

(49) Lammers, S.; Meuwly, M. On the Relationship between Infrared Spectra of Shared Protons in Different Chemical Environments: A Comparison of Protonated Diglyme and Protonated Water Dimer. *J. Phys. Chem. A* **2007**, *111*, 1638–1647.

(50) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry Held in Jerusalem, Israel, April 13-16, 1981*; Pullman, B., Ed.; Springer Netherlands: Dordrecht, 1981; pp 331–342.

(51) Heindel, J. P.; Yu, Q.; Bowman, J. M.; Xantheas, S. S. Benchmark Electronic Structure Calculations for H3O+(H2O)(n), n=0-5, Clusters and Tests of an Existing 1,2,3-Body Potential Energy Surface with a New 4-Body Correction. *J. Chem. Theo. Comp.* **2018**, *14*, 4553–4566.

(52) Vendrell, O.; Gatti, F.; Meyer, H.-D. Dynamics and Infrared Spectroscopy of the Protonated Water Dimer. *Angew. Chem. Int. Ed.* **2007**, *46*, 6918–6921.

(53) Hammer, N. I.; Diken, E. G.; Roscioli, J. R.; Johnson, M. A.; Myshakin, E. M.; Jordan, K. D.; McCoy, A. B.; Huang, X.; Bowman, J. M.; Carter, S. The Vibrational Predissociation Spectra of the $H_5O_2 \cdot RG_n(RG = Ar, Ne)$ Clusters: Correlation of the

Solvent Perturbations in the Free OH and Shared Proton Transitions of the Zundel Ion. *J. Chem. Phys.* **2005**, *122*, 244301.

(54) Fournier, J. A.; Wolke, C. T.; Johnson, C. J.; Johnson, M. A.; Heine, N.; Gewinner, S.; Schöllkopf, W.; Esser, T. K.; Fagiani, M. R.; Knorke, H. et al. Site-Specific Vibrational Spectral Signatures of Water Molecules in the Magic $H_3O^+(H_2O)_{20}$ and $Cs^+(H_2O)_{20}$ Clusters. *Proc. Natl. Acad. Sci.* **2014**, *111*, 18132–18137.

(55) Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. Spectral Signatures of Hydrated Proton Vibrations in Water Clusters. *Science* **2005**, *308*, 1765–1769.

(56) Lee, M. W.; Carr, J. K.; Göllner, M.; Hamm, P.; Meuwly, M. 2D IR Spectra of Cyanide in Water Investigated by Molecular Dynamics Simulations. *J. Chem. Phys.* **2013**, *139*, 054506.

(57) Sidler, D.; Meuwly, M.; Hamm, P. An Efficient Water Force Field Calibrated Against Intermolecular THz and Raman Spectra. *J. Chem. Phys.* **2018**, *148*, 244504.

(58) Yang, N.; Duong, C. H.; Kelleher, P. J.; McCoy, A. B.; Johnson, M. A. Deconstructing water's diffuse OH stretching vibrational spectrum with cold clusters. *Science* **2019**, *364*, 275–278.

(59) Hunter, K. M.; Shakib, F. A.; Paesani, F. Disentangling Coupling Effects in the Infrared Spectra of Liquid Water. *J. Phys. Chem. B* **2018**, *122*, 10754–10761.

(60) Kananenka, A. A.; Skinner, J. L. Fermi resonance in OH-stretch vibrational spectroscopy of liquid water and the water hexamer. *J. Chem. Phys.* **2018**, *148*, 244107.

(61) Nutt, D.; Meuwly, M. Theoretical investigation of infrared spectra and pocket dynamics of photodissociated carbonmonoxy myoglobin. *Biophys. J.* **2003**, *85*, 3612–3623.

(62) Nutt, D.; Meuwly, M. CO migration in native and mutant myoglobin: Atomistic simulations for the understanding of protein function. *Proc. Natl. Acad. Sci.* **2004**, *101*, 5998–6002.

(63) Plattner, N.; Meuwly, M. The role of higher CO-multipole moments in understanding the dynamics of photodissociated carbonmonoxide in myoglobin. *Biophys. J.* **2008**, *94*, 2505–2515.

(64) Cazade, P. A.; Bereau, T.; Meuwly, M. Computational two-dimensional infrared spectroscopy without maps: N-methylacetamide in water. *J. Phys. Chem. B* **2014**, *in print*.

(65) Huang, S.; Braams, B. J.; Bowman, J. M. Ab Initio Potential Energy and Dipole Moment Surfaces for $H_5O_2^+$. *J. Chem. Phys.* **2005**, *122*, 044308.

(66) Thaemer, M.; De Marco, L.; Ramasesha, K.; Mandal, A.; Tokmakoff, A. Ultrafast 2D IR Spectroscopy of the Excess Proton in Liquid Water. *Science* **2015**, *350*, 78–82.

(67) Daly, C. A., Jr.; Streacker, L. M.; Sun, Y.; Pattenaude, S. R.; Hassanali, A. A.; Petersen, P. B.; Corcelli, S. A.; Ben-Amotz, 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*, 5246–5252.

(68) 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.* **1995**, *99*, 5749–5752.

(69) Biswas, R.; Tse, Y.-L. S.; Tokmakoff, A.; Voth, G. A. Role of Presolvation and Anharmonicity in Aqueous Phase Hydrated Proton Solvation and Transport. *J. Phys. Chem. B* **2016**, *120*, 1793–1804.

(70) Goyal, P.; Elstner, M.; Cui, Q. Application of the SCC-DFTB Method to Neutral and Protonated Water Clusters and Bulk Water. *J. Phys. Chem. B* **2011**, *115*, 6790–6805.

(71) Maupin, C. M.; Aradi, B.; Voth, G. A. The Self-Consistent Charge Density Functional Tight Binding Method Applied to Liquid Water and the Hydrated Excess Proton: Benchmark Simulations. *J. Phys. Chem. B* **2010**, *114*, 6922–6931.

(72) Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K. Ions in Water: The Microscopic Structure of a Concentrated HCl Solution. *J. Chem. Phys.* **2004**, *121*, 7840–7848.

(73) Soper, A. K.; Benmore, C. J. Quantum Differences between Heavy and Light Water. *Phys. Rev. Lett.* **2008**, *101*, 065502.

(74) Knight, C.; Lindberg, G. E.; Voth, G. A. Multiscale reactive molecular dynamics. *J. Chem. Phys.* **2012**, *137*, 22A525.

(75) Köfinger, 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*, 2349–2356.

(76) Kumagai, N.; Kawamura, K.; Yokokawa, T. An Interatomic Potential Model for H_2O: Applications to Water and Ice Polymorphs. *Mol. Simul.* **1993**, *12*, 177–186.

(77) Plattner, N.; Meuwly, M. Atomistic Simulations of Co Vibrations in Ices Relevant to Astrochemistry. *Chem. Phys. Chem.* **2008**, *9*, 1271–1277.

(78) Lobaugh, J.; Voth, G. The Quantum Dynamics of an Excess Proton in Water. *J. Chem. Phys.* **1996**, *104*, 2056–2069.

(79) Mei, H.; Tuckerman, M.; Sagnella, D.; Klein, M. Quantum Nuclear Ab Initio Molecular Dynamics Study of Water Wires. *J. Phys. Chem. B* **1998**, *102*, 10446–10458.

(80) Rossi, M.; Ceriotti, M.; Manolopoulos, D. E. Nuclear Quantum Effects in H+ and OH- Diffusion along Confined Water Wires. *J. Phys. Chem. Lett.* **2016**, *7*, 3001–3007.

(81) Marx, D.; Chandra, A.; Tuckerman, M. E. Aqueous Basic Solutions: Hydroxide Solvation, Structural Diffusion, and Comparison to the Hydrated Proton. *Chem. Rev.* **2010**, *110*, 2174–2216.

(82) Kamerlin, S. C. L.; Cao, J.; Rosta, E.; Warshel, A. On Unjustifiably Misrepresenting the EVB Approach While Simultaneously Adopting It. *J. Phys. Chem. B* **2009**, *113*, 10905–10915.

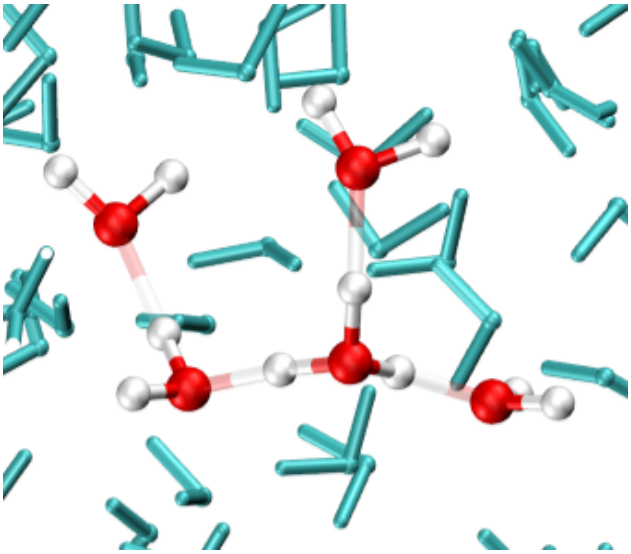
(83) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(84) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theo. Comp.* **2019**, *15*, 3678–3693.

(85) Rivero, U.; Unke, T., Oliver; Meuwly, M.; Willitsch, S. A computational study of the Diels-Alder reactions between 2,3-dibromo-1,3-butadiene and maleic anhydride. *J. Chem. Phys.* **2019**, *151*, 104301.

(86) Pan, S. J.; Yang, Q. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* **2009**, *22*, 1345–1359.

(87) Carpenter, W. B.; Lewis, N. H. C.; Fournier, J. A.; Tokmakoff, A. Entropic barriers in the kinetics of aqueous proton transfer. *J. Chem. Phys.* **2019**, *151*, 034501.



TOC Graphics: A snapshot of a bulk simulation using the MS-MMPT method; different states lead to different H-bonding patterns (transparent OH bonds) for a given conformation.