

The determination of free energy of hydration of water ions from first principles.

Oleg Butin¹, Leonid Pereyaslavets^{1*}, Ganesh Kamath¹, Alexey Illarionov¹, Serzhan Sakipov¹, Igor V Kurnikov¹, Ekaterina Voronina^{1,2}, Ilya Ivahnenko¹, Igor Leontyev¹, Grzegorz Nawrocki¹, Mikhail Darkhovskiy¹, Michael Olevanov^{1,3}, Yevhen K. Cherniavskiy¹, Christopher Lock^{1,4}, Sean Greenslade¹, Roger D Kornberg⁵, Michael Levitt⁵, and Boris Fain^{1*}

¹ InterX Inc. (a subsidiary of NeoTX Therapeutics LTD), 805 Allston Way, Berkeley, CA, 94710, USA

² Lomonosov MSU, Skobeltsyn Institute of Nuclear Physics, Moscow 119991, Russia

³ Lomonosov MSU, Dept. of Physics, Moscow 119991, Russia

⁴ Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Palo Alto, CA 94304, USA

⁵ Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94304, USA

* Corresponding authors: boris.fain@interxinc.com, leonid.pereyaslavets@interxinc.com

Abstract

We model the auto-ionization of water by determining the free energy of hydration of the major intermediate species of water ions. We represent the smallest ions - the hydroxide ion OH^- , the hydronium ion H_3O^+ , and the Zundel ion H_5O_2^+ , - by bonded models; and the more extended ionic structures by strong non-bonded interactions (e.g. the Eigen $\text{H}_9\text{O}_4^+ = \text{H}_3\text{O}^+ + 3(\text{H}_2\text{O})$ and the Stoyanov $\text{H}_{13}\text{O}_6^+ = \text{H}_5\text{O}_2^+ + 4(\text{H}_2\text{O})$). Our models are faithful to the precise QM energies and their components to within 1% or less. Using the calculated free energies and atomization energies, we compute the pK_a of pure water from first principles as a consistency check, and arrive at a value within 1.3 log units of the experimental one. From these calculations we conclude that the hydronium ion, and its hydrated state, the Eigen cation, are the dominant species in the water auto-ionization process.

Introduction

Water auto-ionization is a fundamental chemical process that is enormously important in biology because water ionic species (hydronium and hydroxide ions) are involved in many critical reactions in the organism. For example, the pH of blood in humans needs to be maintained

between 7.35 and 7.45 for metabolic processes to proceed normally¹, with deviations of more than 0.4 pH units away from this range resulting in death. Correct determination of the energetics of the intermediate and final states of water ionization and of other acid-base processes by molecular simulations will provide critical insights into countless biochemical systems.

A quantitative description of the water dissociation equilibrium ($2 \text{H}_2\text{O} \rightleftharpoons \text{OH}^- + \text{H}_3\text{O}^+$) is very challenging, because the hydration free energies of the hydronium and hydroxide species that enter the expression for the free energy of this reaction (Eq 1) are on the order of ~100 kcal/mol, and, thus, should be accurate to within 1% to achieve chemical accuracy.

$$\Delta G_{2 \text{H}_2\text{O} \rightarrow \text{H}_3\text{O}^+ + \text{OH}^-}(\text{aq}) = \Delta G_{2 \text{H}_2\text{O} \rightarrow \text{H}_3\text{O}^+ + \text{OH}^-}(\text{g}) + \Delta G_{\text{hydr}}(\text{H}_3\text{O}^+) + \Delta G_{\text{hydr}}(\text{OH}^-) \quad (\text{eq 1})$$

There are also large uncertainties in the accepted measured values for $\Delta G_{\text{hydr}}(\text{H}^+)$ and $\Delta G_{\text{hydr}}(\text{OH}^-)$ on the experimental side²⁻⁶. These translate into uncertainties of hydration free energy values for other ionic species that are typically determined in experimental setups as relative quantities vs $\Delta G_{\text{hydr}}(\text{H}^+)$.

Researchers have deployed multiple theoretical methods to compute the pK_a values for water. One of the most direct approaches is the *ab initio* molecular dynamics (AIMD^{7,8}) method. In AIMD the potential energy surface of the system is modeled using computational quantum chemical methods (i.e. density functional theory or DFT⁹⁻¹¹), and the atom coordinates are propagated using either classical MD or path integral molecular dynamics (PIMD). Because a system modeling bulk chemical processes such as water dissociation or ion solvation should contain hundreds of atoms, linear scaling DFT methods are typically used¹². There are consistent and ongoing efforts to improve DFT functionals, and while earlier calculations were done in local density approximation (LDA) and generalized gradient approximation (GGA)¹³, more recent AIMD calculations have used meta-GGA functionals with dispersion corrections¹⁴ which produce a more faithful representation of intermolecular interactions. Interaction energies between ions and water are extremely strong, and consequently the accuracy provided by DFT methods may be not sufficient to describe water accurately. Coupled cluster methods (CCM)^{15,16} provide much more reliable estimates of intermolecular interaction energies than DFT, but these scale very poorly with the size of the molecular system, making them unsuitable for bulk AIMD simulations. The cluster continuum solvation (CCM) model^{17,18} and a QM/MM approach

^{19,20} reduce expensive quantum chemical calculations to a small cluster of water molecules interacting with a proton, while bulk water molecules are considered in a continuum or a molecular mechanics approximation. In ²¹ the authors computed proton water clusters on CCSD(T)/CBS level containing up to 6 water molecules. Long-range solvation effects were taken into account by a polarizable continuum approach, while enthalpy and entropy of solvation were computed using the harmonic solvation model. Other groups have deployed less accurate methods for cluster QM calculations that allowed them to consider larger proton-water complexes and different approximations for long-range interactions and solvation free energy estimates. Some of the methods also employ empirical adjustments (e.g. Newtonian energy and force terms) to better reproduce various structural experimental data ²² such as the RDF. The empirical adjustments are made to better reproduce structural information ^{22,23} and may not be 'pressing the correct levels' energetically. In summary, computed values for proton solvation free energy vary substantially within ± 5 kcal/mol, with the currently most widely accepted value of ~ -262 kcal/mol (see ²⁴ or Table SI 4), and the experimental estimates vary in a similar range as the underlying assumptions of the experimental techniques are still debated^{3,6}.

Here we model this critical process by determining the free energy of hydration of the major intermediate species of water ions. We represent the smallest ions - the hydroxide ion OH^- , the hydronium ion H_3O^+ and the Zundel ion H_5O_2^+ - by bonded models, and the more extended ionic structures by strong non-bonded interactions (i.e. the Eigen $\text{H}_9\text{O}_4^+ = \text{H}_3\text{O}^+ + 3(\text{H}_2\text{O})$ and the Stoyanov $\text{H}_{13}\text{O}_6^+ = \text{H}_2\text{O}_2^+ + 4(\text{H}_2\text{O})$, and so on). Our models do not include the ultrafast transitions between the ionic QM minima ^{22,25}, but rather represent the process via the energetics of a series of stable longer-lived intermediates. This is, admittedly, an approximation and interconversion of various states will introduce a small entropic contribution to the final answer. However, as seen from many investigations of the QM energies and interconversions of the states (see detailed review of literature in Table SI 1), the major species we consider correspond to the two main local energy minima with significant barriers between them in the gas phase. It is likely that this separation is preserved in water at room temperature and pressure, and we model it with conventional MD/PIMD where bonds are preserved. This assumption is further supported by the preservation throughout all trajectories of the extended Stoyanov and Eigen ions which are allowed to convert to, respectively, the Zundel and the Hydronium ions in our methodology. Clearly this is a limitation of the current work, and the techniques we use are not suitable for valuable investigations of proton diffusion and kinetics in water.

Our approach is based on the fitting of an atomic polarizable force field to accurate quantum chemical calculations ('silver standard'²⁶) on small molecular fragments (monomers and dimers of chemical functional groups)²⁷⁻³⁰. Using ARROW, our *ab-initio* parameterized force field²⁹, we have been able to predict the free energy of solvation of neutral chemical species dissolved in water and other neutral solvents to within 0.3 kcal/mol of experimental values.²⁹ To further increase the FF - QM correspondence for strong interactions, we have recently augmented the ARROW force field with a neural network term³¹ at short range, which resulted in, among other things, accurate energetics of ion solvation (within 1.5% or ~3 kcal/mol from the experimental values, which themselves have a similar confidence interval)³². NN corrections to ARROW also greatly improved the accuracy of calculations of relative binding free energies of ligands in proteins for systems containing ionic interactions³².

Our approach accurately describes inter- and intra-molecular interactions both at short and long intermolecular distances (silver standard²⁶) and computes the free energy of solvation directly through bulk MD or PIMD simulations. As such, it may have advantages over existing methodologies that compromise on the accuracy of the description of intermolecular interactions' and/or of the description of dynamics of the solvated ion system, and are also hindered by assumptions and approximations involved in combining the various near and far methods²¹.

Models

The basic methodology of the ARROW-FF and ARROW-NN force-fields has been described in several publications^{28-30,32} and is provided in SI 1.1, 1.2, 1.3 and 1.4. The procedure for extracting the dimers and monomers for the training/test sets of the ARROW-NN construction is also provided in SI, sections: SI 3, SI 3.1-3.4. The augmentation of bonded energies by neural networks is described in³² and also in SI 3.6-3.7.

Bonded base states and intermolecular interactions

We describe the widely known ionic species OH^- , H_3O^+ and H_5O_2^+ via bonded intramolecular interactions fitted to reproduce the QM distortion energies (See schematic in Figure 1). The agreement between Force-Field and QM energies for each of these is shown in Table SI 2 and Figures SI 3,4, and the MAE of the agreement is good (0.01, 0.144, 0.122 kcal/mol respectively). Following this, the more extended states such as the Eigen and the Stoyanov ions are formed by non-bonded interactions between water and the three ions above (see Figure 1), with the interaction now being fitted to dimer intermolecular QM energies. The intermolecular interaction correspondence between QM and FF energies is shown in Figure SI 2. The MAE for $(\text{H}_2\text{O} : \text{OH}^-)$ $(\text{H}_2\text{O} : \text{H}_3\text{O}^+)$ and $(\text{H}_2\text{O} : \text{H}_5\text{O}_2^+)$ 2-body components of the total energy are 0.056, 0.092, 0.092 kcal/mol, see Table SI 2. The many-body correspondence is checked on trimers of the constituent ions with 2 water molecules in SI 4, and is within 1.2 kcal/mol. Conventional bonded terms were unable to properly describe the H_5O_2^+ ion which prompted us to add a variation of our neural net term to achieve a good agreement with bonded QM energies of the ion.

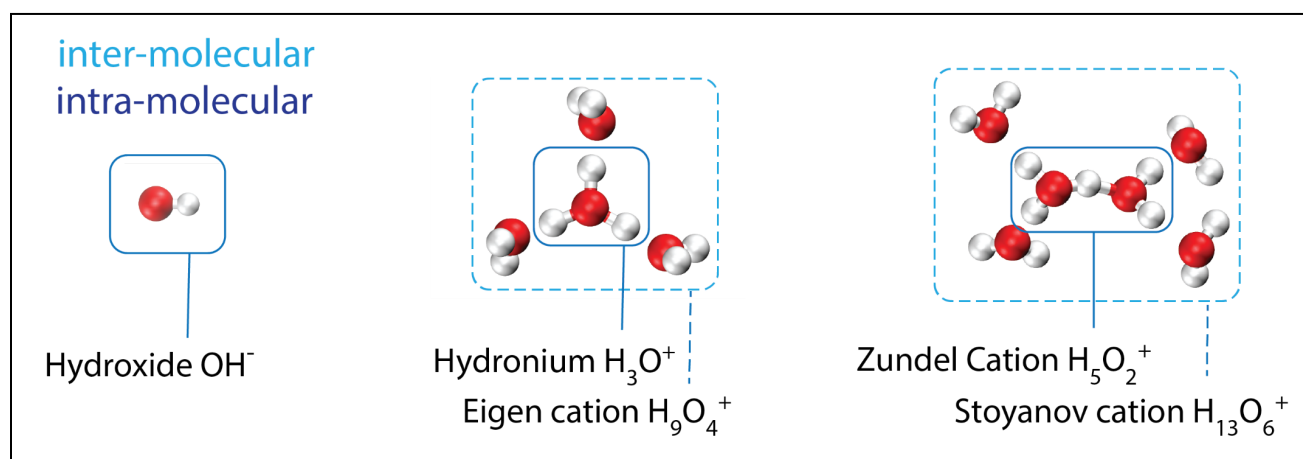


Figure 1: A diagram of the construction of the major anions and cations of water. The hydroxide OH^- , hydronium H_3O^+ and the Zundel cation H_5O_2^+ are modeled via intra-molecular or bonded potentials. The more extended states - the Eigen cation H_9O_4^+ and the Stoyanov cation $\text{H}_{13}\text{O}_6^+$ - are modeled as H_2O molecules with strong bonds to the previous smaller ions.

Simulation and results

We first validated the inter/intramolecular classification of ionic intermediates and species. During the course of the simulation (1 ns) the three waters flanking the hydronium ion remained

attached and unpermuted, as did the four water molecules immediately surrounding the bonded Zundel ion. This was expected as the dimerization energy for $\text{H}_3\text{O}^+ : \text{H}_2\text{O}$ is -51.79 kcal/mol and for $\text{H}_5\text{O}_2^+ : \text{H}_2\text{O}$ is -26.85 kcal/mol. This supports the validity of our decomposition of the more extended ions as the base ions with H_2O molecules attached via QM fitted strong non-bonded interactions, (see Figure 3).

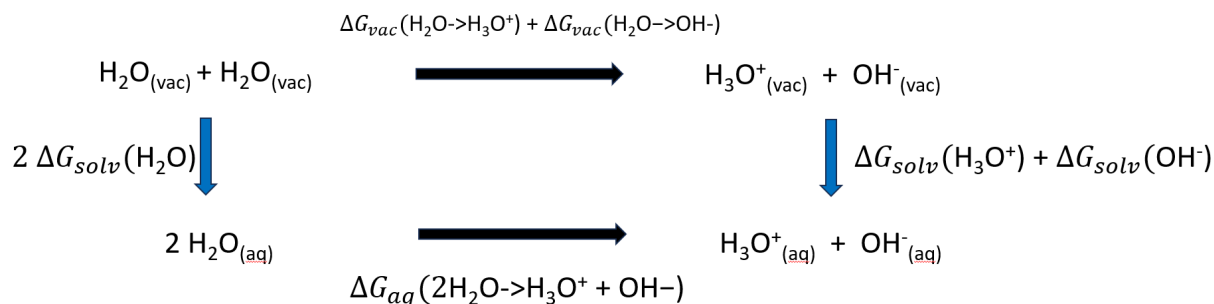


Figure 2. Thermodynamic cycle for the pK_a of water via vacuum alchemical transformations of H_3O^+ and OH^- ions to water and solvation free energies of H_3O^+ and OH^- .

The thermodynamic equilibrium for the pK_a of water is $[\text{H}_2\text{O}]_{(\text{aq})} \leftrightarrow [\text{H}^+]_{(\text{aq})} + [\text{OH}^-]_{(\text{aq})}$, where (aq) stands for the process taking place in liquid water. The naked proton $[\text{H}^+]_{(\text{aq})}$, does not exist for long in the presence of multiple surrounding water molecules in solution. The QM energy of $[\text{H}^+] + [\text{H}_2\text{O}]$ is -76.37804944 AU, while that of the $[\text{H}_3\text{O}^+]$ complex is -76.65125267 AU, thus favoring the bound proton over the naked $[\text{H}^+]_{(\text{aq})}$ state by ~ 171 kcal/mol. Consequently, in solution, by a vast predominance, the proton exists in the bound ground states of hydronium, H_3O^+ , the Zundel, H_5O_2^+ , and other more extended complexes involving these ground states. By definition, the relation³³

$$\Delta G(\text{pKa}, \text{H}_2\text{O}) = 14 \text{ pKa units} = 79.89 \text{ kJ/mol} = 19.11 \text{ kcal/mol}$$

can be expanded for hydronium as

$$\Delta G(\text{pKa}, \text{H}_2\text{O}) = G_{\text{aq}}(\text{OH}^-) + G_{\text{aq}}(\text{H}_3\text{O}^+) - 2 \cdot G_{\text{aq}}(\text{H}_2\text{O}) \quad (\text{eq 2}),$$

where G is free energy of ionic species or water, including all electronic structure energy, in an aqueous environment (see e.g. formulas 18,22 in³³).

Representation of the QM potential surface by force-field energies requires extra care when molecular formation is involved. While in intermolecular interactions the ground state is clearly set to zero at infinite separation, setting the ground state for the intra-molecular energies requires some additional care. These energies represent the electron-electron, electron-nuclei and nuclei-nuclei interactions. Force fields typically set the ground state of molecules (in our case those are H_2O , H_3O^+ , $H_5O_2^+$, OH^-) to 0 or to a small value depending on how intramolecular terms are constructed and parameterized. But in this case the various molecules we evaluate are related through a chemical reaction and therefore their bonded ground states need to match. To do this, for each ion or water (i.e. molecule) we synch up the force field and electronic structure energy through recalibration of the FF ground state:

$$G_{aq}(Molecule) = G_{gr}^{qm}(Molecule) - G_{gr}^{ff}(Molecule) + G_{aq}^{ff}(Molecule) \quad (\text{eq 3}),$$

where $G_{gr}^{qm}(Molecule)$ is the quantum-mechanical energy in ground state of the respective molecule, $G_{gr}^{ff}(Molecule)$ is force field model energy in ground state of the molecule, and $G_{aq}^{ff}(Molecule)$ is the force-field energy of this molecule and its aqueous environment.

When force-field models determine hydration energies, the solute-solvent interactions are fully decoupled^{29,30}. This implicitly includes the alchemical transition to the gas phase,

$$\Delta G_{aq-gas}^{ff} = G_{aq}^{ff} - G_{gas}^{ff}. \text{ The } pK_a \text{ calculation here, however, takes place fully in solution, and}$$

consequently, if we are determining free energies through hydration calculations, mutations in the gas phase need to be explicitly considered. The thermodynamic cycle is shown in Figure 2. Explicitly, $\Delta G(pKa)$ is decomposed into 4 different parts:

$$\begin{aligned} \Delta G(pKa, H_2O) &= \Delta G_{gr}^{qm}(OH^- + H_3O^+ - 2 \cdot H_2O) - \Delta G_{gr}^{ff}(OH^- + H_3O^+ - 2 \cdot H_2O) + \\ &\Delta G_{aq-gas}^{ff}(OH^- + H_3O^+ - 2 \cdot H_2O) + \Delta G_{gas}^{ff}(OH^- + H_3O^+ - 2 \cdot H_2O) \end{aligned} \quad (\text{eq 4}),$$

where

$$\Delta G_{gr}^{qm}(OH^- + H_3O^+ - 2 \cdot H_2O) = G_{gr}^{qm}(OH^-) + G_{gr}^{qm}(H_3O^+) - 2 \cdot G_{gr}^{qm}(H_2O) \quad (\text{eq 5.1})$$

$$\Delta G_{gr}^{ff}(OH^- + H_3O^+ - 2 \cdot H_2O) = G_{gr}^{ff}(OH^-) + G_{gr}^{ff}(H_3O^+) - 2 \cdot G_{gr}^{ff}(H_2O) \quad (\text{eq 5.2})$$

$$\Delta G_{aq-gas}^{ff}(OH^- + H_3O^+ - 2 \cdot H_2O) = \Delta G_{aq-gas}^{ff}(OH^-) + \Delta G_{aq-gas}^{ff}(H_3O^+) - 2 \cdot \Delta G_{aq-gas}^{ff}(H_2O)$$

(eq 5.3)

$$\Delta G_{gas}^{ff}(OH^- + H_3O^+ - 2 \cdot H_2O) = G_{gas}^{ff}(OH^-) + G_{gas}^{ff}(H_3O^+) - 2 \cdot G_{gas}^{ff}(H_2O) \quad (\text{eq 5.4})$$

The formulas for the $H_5O_2^+$ ion are analogous, except that the contributions for 3 and not 2 water molecules need to be subtracted. Similarly, in the gas mutations, $H_5O_2^+$ is transformed into two and not one water molecule. The values of all the above terms calculated with ARROW-NN are summarized in Table 1.

Thermodynamic Free energies G or ΔG (kcal/Mol)	G_{gr}^{qm}	G_{gr}^{ff}	ΔG_{aq-gas}^{ff}	$G_{gas}^{ff}(Ion) - G_{gas}^{ff}(H_2O)$	$G(pK_a, H_2O)^*$
OH^-	-75.74516478	0	-110.07	-2.37	
H_3O^+	-76.65125267	7.52	-101.82	4.37	
$H_5O_2^+$	-153.0819867	18.37	-80.62	24.26	
H_2O	-76.37804944	0	-6.3**	0	
$OH^- + H_3O^+ - 2H_2O$	225.70	7.52	-199.29	2.00	20.89
$OH^- + H_5O_2^+ - 3H_2O$	192.64	18.37	-171.79	21.89	24.37

* The experimental number is 19.11 kcal/mol.

** Here we are using experimental free energy of water desolvation instead of our model's answer. The model self-hydration of H_2O^{29} (-6.15 kcal/mol) is very close to the experimental answer.

Table 1. Thermodynamic data for water and water ions used for estimation of water pK_a values. All units in kcal/mol, except G_{gr}^{qm} which is shown in atomic units (AU) for molecules and kcal/mol for differences.

The hydration energies of OH^- , H_3O^+ and $H_5O_2^+$ via alchemical thermodynamic integration. Full details of the simulation protocols are given in SI 4 and also in previous publications^{29,30}. The last column in Table 1 lists the thermodynamic stability of the various states of water auto-ionization. We see that hydronium, and/or its more complex relative, the Eigen cation, are ~5 kcal/mol more stable than the Zundel cation and/or its relative, the Stoyanov cation, further

corroborated by a recent work in the gas phase performed on a similar QM level³⁴. The H_5O_2^+ is more rarely observed and is likely an important intermediate within the Grotthuss³⁵ mechanism. This observation is supported by recent spectroscopic studies^{36–38} and is a good complement to the excellent theoretical investigation of the dynamics of the system^{14,22,25}.

It is always desirable to connect theoretical studies with a well established observable. In the case of water, the obvious one is its pK_a , 14^{33,39}. The values of Table 1 taken together predict a pK_a of H_3O^+ ion of 15.32 for the PIMD simulations (with classical MD result of 14.98 is in Table SI 3); a satisfyingly accurate result within 1.5 log-units (~ 1.8 kcal/mol) from the experimental number. Figure 3 summarizes our results in a schematic manner. For context, various theoretical estimates of the free energies of solvation for OH^- and H_3O^+ throughout history have a spread of ~ 10 kcal/mol (Table SI 4).

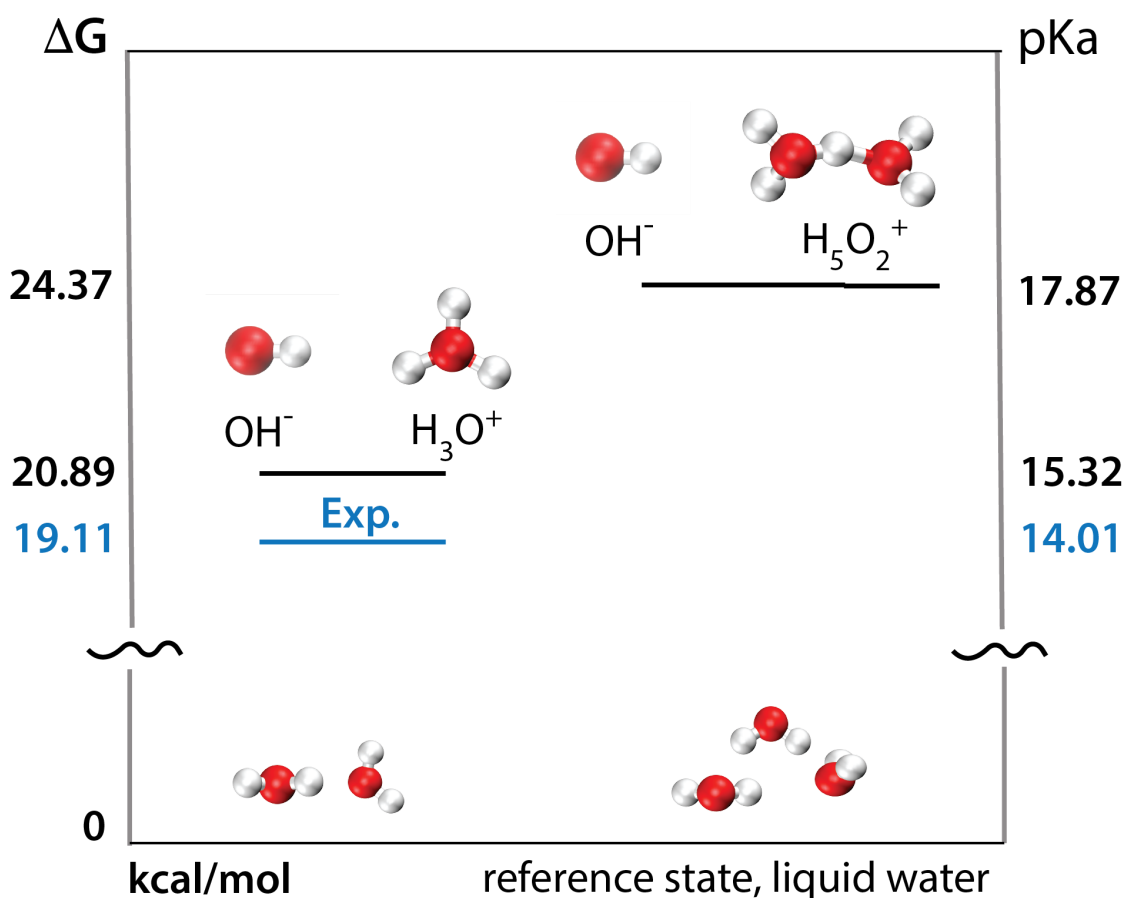


Figure 3: A schematic representation of the $G(pK_a, H_2O)$ levels for the two dominant species of water ions, hydroxide + hydronium and hydroxide + Zundel cation. Our calculations show that hydronium and its extended complexes such as the Eigen cation are the dominant ionic species.

Estimates and sources of errors

Considering the magnitude of the constituent energies (~ 200 kcal/mol), and the complexity and heterogeneity of the system, the fact that our theoretical estimate of pK_a is only 1.8 kcal/mol ($\sim 1\%$) off is satisfying. We also note that in our previous work³¹ the hydration of NaCl and LiCl salts (also ~ 200 kcal/mol) was off by 3 kcal/mol. Some of the sources of error may come from insufficiently precise description of many-body effects, both in our polarization model (i.e. non-linearity of polarization with such high local electric potentials) as well as from the fundamental many-body effects^{40,41} like field-induced dispersion, which our functional forms do not describe. Another source of error may be the fact that we do not describe delocalization and hopping of the proton between water molecules and ions.

Conclusion

Using the ARROW FF-NN approach we computed free energies of the dominant ionized water species. The correctness of the FF-NN hybrid model is underscored by the accurate prediction of water pK_a and by the agreement with experimental results. The ARROW FF-NN methods are a straightforward systematic procedure that does not require any empirical adjustments, and that produces very accurate FF models for complex ionic species of great importance for biocatalysis applications. The construction is computationally tractable as it requires benchmark QM computations consisting of mostly molecular or ionic *dimers*. The same methodology can be applied to chemical reaction intermediates, and will permit computation of free energies of distinct intermediate catalytic states formed during enzymatic catalysis in regular MD simulations. Our methods do not describe the important and intricate proton diffusion and kinetics. However, as a reasonable approximation of quasi-stable dominant states, and one that produces non-adjusted equilibrium prediction of observables, it has significant value.

Acknowledgments

The authors would like to thank Asher Nathan, Robert Harrow, Jeffrey Kletzel, Paul T. Marinelli, and InterX and NeoTX teams for their support.

References

1. Hopkins, E., Sanvictores, T. & Sharma, S. *Physiology, Acid Base Balance*. (StatPearls Publishing, 2022).
2. Marcus, Y. Thermodynamics of solvation of ions. Part 5.—Gibbs free energy of hydration at 298.15 K. *J. Chem. Soc. Faraday Trans.* **87**, 2995–2999 (1991).
3. Tissandier, M. D. *et al.* The Proton's Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data. *J. Phys. Chem. A* **102**, 7787–7794 (1998).
4. Camaioni, D. M. & Schwerdtfeger, C. A. Comment on 'Accurate experimental values for the free energies of hydration of H⁺, OH⁻, and H₃O⁺'. *J. Phys. Chem. A* **109**, 10795–10797 (2005).
5. Fawcett, W. R. The ionic work function and its role in estimating absolute electrode potentials. *Langmuir* **24**, 9868–9875 (2008).
6. Vlcek, L., Chialvo, A. A. & Simonson, J. M. Correspondence between cluster-ion and bulk solution thermodynamic properties: on the validity of the cluster-pair-based approximation. *J. Phys. Chem. A* **117**, 11328–11338 (2013).
7. Iftimie, R., Minary, P. & Tuckerman, M. E. Ab initio molecular dynamics: concepts, recent developments, and future trends. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 6654–6659 (2005).
8. Car, R. & Parrinello, M. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **55**, 2471–2474 (1985).
9. Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
10. Pribram-Jones, A., Gross, D. A. & Burke, K. DFT: A Theory Full of Holes? *Annu. Rev. Phys.*

- Chem.* **66**, 283–304 (2015).
11. Teale, A. M. *et al.* DFT exchange: sharing perspectives on the workhorse of quantum chemistry and materials science. *Phys. Chem. Chem. Phys.* **24**, 28700–28781 (2022).
 12. Mohr, S. *et al.* Accurate and efficient linear scaling DFT calculations with universal applicability. *Phys. Chem. Chem. Phys.* **17**, 31360–31370 (2015).
 13. Trout, B. L. & Parrinello, M. Analysis of the Dissociation of H₂O in Water Using First-Principles Molecular Dynamics. *J. Phys. Chem. B* **103**, 7340–7345 (1999).
 14. Wang, R., Carnevale, V., Klein, M. L. & Borguet, E. First-Principles Calculation of Water pK_a Using the Newly Developed SCAN Functional. *J. Phys. Chem. Lett.* **11**, 54–59 (2020).
 15. Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **45**, 4256–4266 (1966).
 16. Crawford, T. D. & Schaefer, H. F., III. An introduction to coupled cluster theory for computational chemists. in *Reviews in Computational Chemistry* 33–136 (John Wiley & Sons, Inc., 2007).
 17. Tawa, G. J., Topol, I. A., Burt, S. K., Caldwell, R. A. & Rashin, A. A. Calculation of the aqueous solvation free energy of the proton. *J. Chem. Phys.* **109**, 4852–4863 (1998).
 18. Zhan, C.-G. & Dixon, D. A. Absolute Hydration Free Energy of the Proton from First-Principles Electronic Structure Calculations. *J. Phys. Chem. A* **105**, 11534–11540 (2001).
 19. Duignan, T. T., Baer, M. D., Schenter, G. K. & Mundy, C. J. Real single ion solvation free energies with quantum mechanical simulation. *Chem. Sci.* **8**, 6131–6140 (2017).
 20. Hofer, T. S. & Hünenberger, P. H. Absolute proton hydration free energy, surface potential of water, and redox potential of the hydrogen electrode from first principles: QM/MM MD free-energy simulations of sodium and potassium hydration. *J. Chem. Phys.* **148**, 222814 (2018).

21. Ishikawa, A. & Nakai, H. Quantum chemical approach for condensed-phase thermochemistry (III): Accurate evaluation of proton hydration energy and standard hydrogen electrode potential. *Chem. Phys. Lett.* **650**, 159–164 (2016).
22. Calio, P. B., Li, C. & Voth, G. A. Resolving the Structural Debate for the Hydrated Excess Proton in Water. *J. Am. Chem. Soc.* **143**, 18672–18683 (2021).
23. White, A. D., Knight, C., Hocky, G. M. & Voth, G. A. Communication: Improved ab initio molecular dynamics by minimally biasing with experimental data. *J. Chem. Phys.* **146**, 041102 (2017).
24. Ho, J. & Coote, M. L. A universal approach for continuum solvent pKa calculations: are we there yet? *Theor. Chem. Acc.* **125**, 3–21 (2010).
25. Knight, C. & Voth, G. A. The Curious Case of the Hydrated Proton. *Acc. Chem. Res.* **45**, 101–109 (2012).
26. Burns, L. A., Marshall, M. S. & Sherrill, C. D. Appointing silver and bronze standards for noncovalent interactions: a comparison of spin-component-scaled (SCS), explicitly correlated (F12), and specialized wavefunction approaches. *J. Chem. Phys.* **141**, 234111 (2014).
27. Donchev, A. G. *et al.* Assessment of performance of the general purpose polarizable force field QMPFF3 in condensed phase. *J. Comput. Chem.* **29**, 1242–1249 (2008).
28. Kamath, G. *et al.* Prediction of cyclohexane-water distribution coefficient for SAMPL5 drug-like compounds with the QMPFF3 and ARROW polarizable force fields. *J. Comput. Aided Mol. Des.* **30**, 977–988 (2016).
29. Pereyaslavets, L. *et al.* Accurate determination of solvation free energies of neutral organic compounds from first principles. *Nat. Commun.* **13**, 414 (2022).
30. Nawrocki, G. *et al.* Protein-ligand binding free-energy calculations with ARROW—A purely first-principles parameterized polarizable force field. *J. Chem. Theory Comput.* **18**, 7751–7763 (2022).

31. Alexey Illarionov, Serzhan Sakipov, Leonid Pereyaslavets, Igor V. Kurnikov, Ganesh Kamath, Oleg Butin, Ekaterina Voronina, Ilya Ivahnenko, Igor Leontyev, Grzegorz Nawrocki, Mikhail Darkhovskiy, Michael Olevanov, Yevhen K. Cherniavskiy, Christopher Lock, Sean Greenslade, Subramanian Sankaranarayanan, Maria Kurnikova, Jeffrey Potoff, Roger D. Kornberg, Michael Levitt, and Boris Fain. Combining Force Fields and Neural Networks for an Accurate Representation of Chemically Diverse Molecular Interactions. *J. Am. Chem. Soc.* **145**, 23620–23629 (2023).
32. Illarionov, A. *et al.* Fast and Accurate Representation of Diverse Molecular Interactions. *Submitted to J. Am. Chem. Soc.* (2023).
33. Meister, E. C., Willeke, M., Angst, W., Togni, A. & Walde, P. Confusing quantitative descriptions of Brønsted–Lowry Acid–Base equilibria in chemistry textbooks - A critical review and clarifications for chemical educators. *Helv. Chim. Acta* **97**, 1–31 (2014).
34. Schran, C., Behler, J. & Marx, D. Automated Fitting of Neural Network Potentials at Coupled Cluster Accuracy: Protonated Water Clusters as Testing Ground. *J. Chem. Theory Comput.* **16**, 88–99 (2020).
35. de Grotthuss, C. J. T. Sur la décomposition de l'eau et des corps qu'elle tient en dissolution à l'aide de l'électricité galvanique. *Ann. Chim.* **58**, 54–73 (1806).
36. Stoyanov, E. S., Stoyanova, I. V. & Reed, C. A. The structure of the hydrogen ion (H(aq)⁺) in water. *J. Am. Chem. Soc.* **132**, 1484–1485 (2010).
37. Kulig, W. & Agmon, N. Both Zundel and Eigen isomers contribute to the IR spectrum of the gas-phase H₉O₄⁺ cluster. *J. Phys. Chem. B* **118**, 278–286 (2014).
38. Schröder, M., Gatti, F., Lauvergnat, D., Meyer, H.-D. & Vendrell, O. The coupling of the hydrated proton to its first solvation shell. *Nat. Commun.* **13**, 6170 (2022).
39. Silverstein, T. P. & Heller, S. T. pK_a Values in the Undergraduate Curriculum: What Is the Real pK_a of Water? *J. Chem. Educ.* **94**, 690–695 (2017).
40. Kleshchonok, A. & Tkatchenko, A. Tailoring van der Waals dispersion interactions with

external electric charges. *Nat. Commun.* **9**, 3017 (2018).

41. Ambrosetti, A., Ferri, N., DiStasio, R. A., Jr & Tkatchenko, A. Wavelike charge density fluctuations and van der Waals interactions at the nanoscale. *Science* **351**, 1171–1176 (2016).