Acids at the edge: why nitric and formic acid dissociations at air-water interfaces depend on depth and on interface specific area

Miguel de la Puente, Rolf David, Axel Gomez, Damien Laage

Acids at the edge: why nitric and formic acid dissociations at air-water interfaces depend on depth and on interface specific area

Miguel de la Puente, Rolf David, Axel Gomez, and Damien Laage*

PASTEUR, Department of Chemistry, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France

E-mail: damien.laage@ens.psl.eu

Abstract

Whether the air-water interface weakens or strengthens the acidity of simple organic and inorganic acids compared to the bulk has a critical importance in a broad range of environmental and biochemical processes. However, consensus has not yet been achieved on this key question. Here we use machine learning-based reactive molecular dynamics simulations to study the dissociation of the paradigmatic nitric and formic acids at the air-water interface. We show that the local acidity profile across the interface is determined by changes in acid and conjugate base solvation and that acidity drops abruptly over a transition region of a few molecular layers. At the interface, both acids are weaker than in the bulk due to desolvation. In contrast, below the interface acidities reach a plateau and are all the stronger compared to the bulk as the aqueous phase surface/volume ratio is large, due to the growing impact of the released proton stabilization at the water surface. These results imply that the measured degree of dissociation sensitively depends on the experimental probing length and system size, and

suggest a molecular explanation to the contrasting experimental results. The aerosol size dependence of acidity has important consequences for atmospheric chemistry.

Introduction

Air-water interfaces are ubiquitous and central to many essential chemical, biochemical and environmental processes, ranging from atmospheric reactions on aerosols to ocean acidification and chemical exchanges in the lungs.[1–3] There is now growing evidence that chemical reactivity at air-water interfaces can be radically different from that in the bulk,[3–5] and some reactions that are slow or unfavorable in the bulk have been observed[3] to proceed much more easily at the air-water interface. This provides promising green chemistry routes for chemical synthesis[4] and could have played an important role in the origins of life, e.g., by catalyzing the prebiotic assembling of peptides[6] and nucleic acids.[7]

However, while the air-water interface facilitates some reactions compared to the bulk, it impedes others, and an explanation for these contrasting effects has remained elusive. The paradigm case of acid dissociations at the air-water interface is particularly revealing. Extensive experimental[8–17] and simulation[11,18–22] studies have characterized the change in acidity between bulk solution and air-water interface for a series of organic and inorganic acids. Both acidity enhancements[16,17,21,22] and reductions[9–13,15,18–20] have been reported, sometimes for the same acid, thus suggesting that interfacial effects on acidity depend on the nature of the acid, but also possibly on the interfacial region that is probed.

Molecular dynamics simulations are an exquisite tool to identify the molecular factors that cause a change in the acid-base equilibrium between bulk and interface. However, the covalent bond breaking and forming occurring during acid dissociation preclude the use of inexpensive classical force-field simulations and require electronic structure calculations. The latter are typically performed via DFT-based molecular dynamics simulations,[11,18–22] but their large computational cost imposes strong limitations on the trajectory lengths, and

thus on the precision of the computed acid-base free energy difference. However, recent developments in machine learning[23,24] now offer a solution to this long-standing difficulty: deep neural network potentials (NNPs) can be used as reactive force fields with a DFT-level quality at a fraction of the computational cost and have been successfully applied to different contexts including, e.g., bulk aqueous solutions,[25] aqueous interfaces[26] and gas-phase combustion.[27]

Here, we focus on the dissociation of nitric and formic acids and we contrast their acidities in bulk aqueous solution and at different distances above and below the air-water interface. These two acids play very important roles in atmospheric chemistry,[28] and respectively represent typical strong inorganic and weak carboxylic acids. Prior experimental and simulation works on these and closely related acids have led to diverging conclusions about the effect of an air-water interface on their acidities: some studies[10–13,19,20] suggested that dissociation is reduced at the air-water interface while others[16,22] concluded that it is enhanced or accelerated. To elucidate the cause of these interfacial effects, we trained a neural-network-based reactive force field for each acid on reference DFT calculations and we probed the local acidity by determining the degree of dissociation, calculated from the acid dissociation free energy profiles in bulk solution and at the air-water interface. Our results reveal that, in contrast to the intuitive expectation, the air-water interface region does not provide an environment intermediate between these of the bulk liquid and of the vapor phase. Instead, we show that acidity changes abruptly over a few molecular layers when the interface is crossed: for both acids, acidity decreases from a value stronger than in the bulk below the interface to a value smaller than in the bulk above the interface. We present a model that quantitatively connects the local acidity to solvation properties for the acid, conjugate base and released hydronium, and show that the measured acidity sensitively depends on the experimental probing length and on the system specific area.

Methods

We first briefly summarize our NNP methodology (see SI for further details). NNPs are trained with the DeePMD-kit[23] on reference BLYP-D3 energy and force calculations obtained for systems comprising one acid and 128 water molecules. Training set configurations are taken at different stages of the dissociation reaction between the reactant and product states, both in the bulk and at the interface, from a combination of unconstrained trajectories and biased sampling, e.g. via metadynamics, along the dissociation reaction coordinate chosen to be the change in hydrogen coordination number around the acid oxygen atoms. The training set is iteratively enriched to yield a converged NNP for each acid that can describe the dissociation reaction both in the bulk and at the air-water interface at the BLYP-D3 level, for a computational cost that is more than 1,000-fold reduced.

We use these NNPs to determine acid dissociation free energy profiles in the bulk and at different distances from the air-water interface. These profiles are reconstructed from a series of biased trajectories where the acid position in the slab is constrained and the acid is progressively dissociated (see SI). Long ($\simeq$ ns) simulation times are required for each profile to reach the necessary precision on the dissociation free energy (here 0.1 kcal/mol) and would have been unattainable with DFT-based molecular dynamics.

Results and Discussion

In bulk aqueous solution, the calculated dissociation free energy profiles (black curves in Fig. 1) reproduce the strong character of nitric acid and the weak acidity of formic acid. The computed pK_A are in very good agreement with experimental values: pK_A^{calc}=-2.0$\pm$0.1 vs. pK_A^{exp}=-1.4[29] for nitric acid, and pK_A^{calc}=3.5$\pm$0.1 vs. pK_A^{exp}=3.7[29] for formic acid (see Tables S1-S2).

In contrast, when the acids are in the vapor phase on top of the water surface (orange and blue curves in Fig. 1), their dissociation is much less favorable, due to the greater desta-

bilization by the apolar vapor environment of the charged dissociation products (conjugate base anion and hydronium cation) than of the neutral acid form. The acidities of nitric and formic acids are thus strongly reduced and both acids are essentially undissociated.

In the vicinity of the interface, the behavior is more complex and as previously observed for other reactions[30] it contrasts with the simple expectation that interfacial acidity should be intermediate between the bulk and vapor phase limits. Below the interface, dissociation free energies are more favorable than in the bulk, indicating enhanced acidities (green and brown curves); in contrast, acidities are decreased immediately above the interface (purple curves). This change in dissociation free energies takes place abruptly over a few molecular layers.

We now show that these intriguing acidity changes at the air-water interface can be explained by examining how the reaction's reactants and products are solvated at the interface. Figure 2 presents the hydration free energy profiles when the acid, its conjugate base and the hydronium ion are separately moved across the water slab perpendicularly to the interface.

We first consider the situation in the second layer below the interface. For nitric and formic acids, the acid and its conjugate base are fully solvated, as in the bulk. However, the hydronium ion released upon dissociation is free to diffuse to the interface, where it is preferentially solvated with respect to the bulk, which shifts the dissociation equilibrium from its bulk value in favor of the deprotonated form. Our calculations indicate that hydronium is stabilized by approximately 1.0±0.2 kcal/mol with respect to the slab interior (Fig. 2), in good agreement with the 1.3±0.2 kcal/mol value recently determined experimentally.[31] Acidity within the slab is therefore not bulk-like and is enhanced by the presence of the air-water interface. As shown below, this acidity enhancement is increasingly pronounced for aqueous systems with increasing surface/volume ratios and this important result should apply to all acids.

We now turn to acids lying at the interface. Both nitric and formic acids are preferentially solvated with respect to the slab interior and their most stable position is on top of the first

water layer (Fig. 2). Nitric acid tends to lie flat at the interface (see Fig. S4) as previously suggested;[9,18] formic acid points its polar carboxylic acid moiety toward the liquid and its apolar hydrocarbon part toward the vapor, in agreement with recent experiments[32] (see Fig. S5). Incomplete solvation of the negative charge on the nitrate and formate anions at the interface implies that conjugate bases are less well solvated than neutral acid forms. This effect is particularly pronounced for the nitrate ion which is destabilized at the interface, in agreement with its bulk preference found in prior simulations[33,34] and experiments,[35,36] and less so for the amphiphilic formate anion. This base destabilization more than compensates for the favorable hydronium hydration at the interface and leads to a decreased acidity at the interface with respect to the bulk.

The acid dissociation reaction free energy $\Delta_r G(d)$ at depth d in the slab can be approximated as the sum of the bulk reaction free energy $\Delta_r G_{\text{bulk}}$ and of the ΔG_{hyd} hydration free energy changes relative to the bulk determined separately for the acid (A), base (B), hydronium ion (H) and reactant water (W) molecule,

$$\begin{aligned}\Delta_r G(d) &= \Delta_r G_{\text{bulk}} + \\ &\Delta G^B_{\text{hyd}}(d) + \langle \Delta G^H_{\text{hyd}} \rangle_L - \Delta G^A_{\text{hyd}}(d) - \langle \Delta G^W_{\text{hyd}} \rangle_L \end{aligned} \tag{1}$$

The hydronium and water free energies are averaged over the slab thickness L because they can freely diffuse. The impact of the hydronium preferential solvation at the interface therefore increases when the slab becomes thinner; in contrast, for an infinitely thick slab the $\langle \Delta G^{H,W}_{\text{hyd}} \rangle_L$ free energies asymptotically vanish and the dissociation free energy is exclusively determined by the acid and base solvation free energies, as proposed in prior work[37] (see SI).

This local acidity model provides a continuous description of the acidity changes with the distance to the interface and Fig. 3 shows that its predictions are in excellent agreement with the direct calculations at a series of positions in Fig. 1. (Fig. S3 shows that an equally good agreement between model and simulations is obtained for a system that is twice larger.)

How far below the interface the dissociation equilibrium of an acid is affected has been an important but challenging question for the interpretation of experiments.[11] Our results clearly show that for both acids the transition region around the interface where acidity changes abruptly extends over at most two molecular layers. Below this transition region, acidity reaches a plateau which is distinct from the bulk value because of the possibility for the released proton to be stabilized at the interface, and acidity in the core of the aqueous system increases when the surface/volume ratio increases. Our model presents two important distinguishing features with respect to prior descriptions of acidity at interfaces via two-state mixtures of bulk-like and interfacial environments:[38] it accounts for the broad distribution of local solvation environments and most importantly it correctly predicts the increased acidity below the interface. Since solvation free energies entering in eq. 1 do not require electronic structure calculations, an important methodological implication is that interfacial dissociation free energies should be well approximated via inexpensive classical simulations with polarizable force-fields that reproduce the solutes' surface propensities. This model can be extended to describe the dissociation of positively charged acids (e.g., trimethylammonium[39]), where the combination of greater destabilization of the acid relative to the base and hydronium stabilization at the interface is expected to lead to a strong increase in the dissociated fraction, in agreement with a prior theoretical study.[37] Finally, this model can be easily generalized to different geometries and to include additional solvation contributions coming, e.g., from ions which have been shown[13] to affect interfacial acidity.

We can now use this molecular picture relating interfacial acid dissociation and local solvation to propose a molecular explanation to the contrasting conclusions obtained by different experimental techniques. Vibrational sum-frequency generation (SFG)[9,10] and X-ray photoelectron spectroscopy (XPS)[11,12] experiments on inorganic (including nitric) and carboxylic (including formic) acids measured a decreased acidity at the air-water interface. In contrast, online electrospray mass spectrometry (OESI-MS) studies concluded that acidity decreases[13] for nitric acid but increases[16] for some small carboxylic acids (e.g. acetic and

pyruvic acids).

Our model reveals that the differences in probing depth and sample specific area between these two sets of techniques have a major impact on the measured acid dissociated fraction. For XPS and SFG, the signal mainly originates from the topmost interfacial layer where our calculations show that acidity is decreased with respect to the bulk and the sample surface/volume ratio is very small (15-20 μm-thick liquid jet for XPS[11,12] and macroscopic slab for SFG) so that the interfacial proton stabilization effect is limited. Both factors thus explain that the measured dissociated fraction is smaller than in the bulk. We used our model to calculate the degree of dissociation for a thick slab when uniformly probed over the topmost layer to approximate the SFG result and with a 9 Å-exponential probing dependence[11] to model the (sub-nm) electron inelastic mean free path in XPS (see SI). For nitric acid which is fully dissociated in the bulk, our results yield a substantial undissociated fraction within the topmost layer, in agreement with SFG experiments which could detect[9] undissociated nitric acid, and predict an undissociated fraction of approximately 10% when considering an exponential probing dependence (Fig. S10), in remarkable agreement with the 9% fraction reported by XPS in dilute conditions.[11] Regarding the weak formic acid, our model predicts that the conjugate base fraction is much less than 1% in the topmost layer, which agrees both with SFG which did not find any evidence of the formate ion[10] and with XPS measurements[12] on small carboxylic acids which found that the acid form fraction is greater at the interface than in the bulk (To distinguish between formate and formic acid, SFG assumes that the CH vibrational frequency shift induced by deprotonation is similar in the bulk and at the interface; as shown in Fig. S7, our NNP description reproduces the experimental shift and supports this hypothesis). SFG, XPS and our calculations thus concur that nitric and formic acids are weaker at the interface than in the bulk. While nitric acid remains a strong acid at the interface, the small interfacial undissociated fraction has important consequences for atmospherically-relevant photochemical reactions.[40]

We now turn to OESI-MS which measures the composition of (typically nanometer-

to micrometer-sized) aqueous droplets after they collide with gaseous acid molecules and therefore probes systems with large interfacial specific areas. While it is more sensitive to surface processes than traditional ESI-MS, its suggested[16] nanometric probing depth implies that it reports not only on the topmost interfacial layer but also on layers below the interface. The larger proton stabilization combined with the greater probing depth can thus explain why acidity increases have been measured for some acids.[16] While a quantitative description of OESI-MS results would require considering multiple complex factors including, e.g., droplet charge and evaporation[41] and possible non-equilibrium effects, here we simply use our model to determine the acid/base ratio uniformly averaged over a nanometric depth to show that differences in probing length and sample size with SFG and XPS can already explain the contrasting results. Calculations based on our model (Fig. S9) show that for formic acid, a reduced acid/base ratio, i.e., an enhanced acidity, is expected when aqueous systems with a $\simeq$nm thickness are examined with a $\simeq$nm probing depth, which is consistent with the increased acidity measured by OESI-MS for acetic and pyruvic acids.[16] In contrast, for the more hydrophilic nitrate ion, its destabilization due to the interface extends further below the interface (Fig. 2) where it more than compensates the hydronium ion surface affinity and makes nitric acid weaker (see Fig. 3) leading to increased acid/base ratios for all system sizes and probing depths considered in this work (Fig. S9), in agreement with the decreased acidity reported by OESI-MS measurements.[13]

The change in acidity with the aqueous system specific area revealed by our results has important consequences for chemistry at the surface of and within atmospheric aerosols. For example, our model eq. 1 shows that over the typical nm to 100 μm range of aerosol sizes[28] the hydronium hydration free energy $\langle \Delta_{hyd} G^H \rangle_L$ changes by $\simeq$0.5 kcal/mol, causing a $\simeq$0.4 pK_A shift for all acids in the aerosol core at 300 K, which is further enhanced at lower temperature.

We have so far focused on the dissociation thermodynamic equilibrium constant at the interface. Regarding the dissociation rate constant, prior simulations[21,22] proposed that

dissociation could be accelerated at the interface due to the smaller reorganization energy caused by the lower polarity and reduced number of hydrogen-bonds to be rearranged. Our free energy profiles (Fig. 1) suggest that this effect is limited and that the change in reaction free energy barrier is mostly due to the shift in reactant and product state equilibrium solvation free energies, described by our model.

Concluding Remarks

We have thus shown via reactive NNP molecular dynamics simulations of nitric and formic acids that local acidity at the air-water interface is directly determined by solvation of the acid, conjugate base and hydronium ion. As the interface is crossed, acidity decreases over a few molecular layers from a value weaker than in the bulk on the vapor side to a value stronger than in the bulk below the interface in systems with large specific areas. The impacts of the experimental probing depth and system surface/volume ratio on acidity offer a molecular explanation to the contrasting experimental results. Our model provides an improved molecular comprehension of acid dissociation at interfaces and of its system size dependence, of great importance for the characterization of surface composition and for interfacial reactivity in, e.g., atmospheric chemistry and catalysis within aqueous droplets and nanoscopic channels.

Acknowledgement

This work was supported by PSL OCAV (Idex ANR-10-IDEX-0001-02PSL) and an HPC allocation from GENCI–IDRIS (Grant 2020-101591).

Supporting Information Available

The following files are available free of charge. Methodology for neural network potential training and free energy calculations, molecular orientations at interface, formate/formic acid vibrational spectra, local acidity model.

References

(1) Finlayson-Pitts, B. J. Reactions at Surfaces in the Atmosphere: Integration of Experiments and Theory As Necessary (but Not Necessarily Sufficient) for Predicting the Physical Chemistry of Aerosols. *Phys Chem Chem Phys* **2009**, *11*, 7760.

(2) Griffith, E.; Tuck, A.; Vaida, V. Ocean-Atmosphere Interactions in the Emergence of Complexity in Simple Chemical Systems. *Acc Chem Res* **2012**, *45*, 2106–2113.

(3) Ruiz-Lopez, M. F.; Francisco, J. S.; Martins-Costa, M. T.; Anglada, J. M. Molecular Reactions at Aqueous Interfaces. *Nat Rev Chem* **2020**, *4*, 459–475.

(4) Wei, Z.; Li, Y.; Cooks, R. G.; Yan, X. Accelerated Reaction Kinetics in Microdroplets: Overview and Recent Developments. *Annu Rev Phys Chem* **2020**, *71*, 31–51.

(5) Kusaka, R.; Nihonyanagi, S.; Tahara, T. The Photochemical Reaction of Phenol Becomes Ultrafast at the Air-Water Interface. *Nat Chem* **2021**, *13*, 306–311.

(6) Griffith, E. C.; Vaida, V. In Situ Observation of Peptide Bond Formation at the Water–Air Interface. *Proc Natl Acad Sci* **2012**, *109*, 15697–15701.

(7) Nam, I.; Lee, J. K.; Nam, H. G.; Zare, R. N. Abiotic Production of Sugar Phosphates and Uridine Ribonucleoside in Aqueous Microdroplets. *Proc Natl Acad Sci U S A* **2017**, *114*, 12396–12400.

(8) Yang, H.; Finlayson-Pitts, B. J. Infrared Spectroscopic Studies of Binary Solutions of Nitric Acid and Water and Ternary Solutions of Nitric Acid, Sulfuric Acid, and Water

at Room Temperature: Evidence for Molecular Nitric Acid at the Surface. *J Phys Chem A* **2001**, *105*, 1890–1896.

(9) Kido Soule, M.; Blower, P.; Richmond, G. Nonlinear Vibrational Spectroscopic Studies of the Adsorption and Speciation of Nitric Acid at the Vapor/acid Solution Interface. *J Phys Chem A* **2007**, *111*, 3349–3357.

(10) Johnson, C. M.; Tyrode, E.; Kumpulainen, A.; Leygraf, C. Vibrational Sum Frequency Spectroscopy Study of the Liquid/Vapor Interface of Formic Acid/Water Solutions. *J Phys Chem C* **2009**, *113*, 13209–13218.

(11) Lewis, T.; Winter, B.; Stern, A. C.; Baer, M. D.; Mundy, C. J.; Tobias, D. J.; Hemminger, J. C. Does Nitric Acid Dissociate at the Aqueous Solution Surface? *J Phys Chem C* **2011**, *115*, 21183–21190.

(12) Ottosson, N.; Wernersson, E.; Söderström, J.; Pokapanich, W.; Kaufmann, S.; Svensson, S.; Persson, I.; Ohrwall, G.; Björneholm, O. The Protonation State of Small Carboxylic Acids at the Water Surface from Photoelectron Spectroscopy. *Phys Chem Chem Phys* **2011**, *13*, 12261–12267.

(13) Mishra, H.; Enami, S.; Nielsen, R. J.; Hoffmann, M. R.; Goddard, W. A.; Colussi, A. J. Anions Dramatically Enhance Proton Transfer Through Aqueous Interfaces. *Proc Natl Acad Sci* **2012**, *109*, 10228–10232.

(14) Mishra, H.; Enami, S.; Nielsen, R.; Stewart, L.; Hoffmann, M.; Goddard, W.; Colussi, A. Brønsted Basicity of the Air-Water Interface. *Proc Natl Acad Sci* **2012**, *109*, 18679–18683.

(15) Wellen, B.; Lach, E.; Allen, H. Surface PKa of Octanoic, Nonanoic, and Decanoic Fatty Acids at the Air-Water Interface: Applications to Atmospheric Aerosol Chemistry. *Phys Chem Chem Phys* **2017**, *19*, 26551–26558.

(16) Eugene, A.; Pillar-Little, E.; Colussi, A.; Guzman, M. Enhanced Acidity of Acetic and Pyruvic Acids on the Surface of Water. *Langmuir* **2018**, *34*, 9307–9313.

(17) Biswas, B.; Singh, P. C. The Enhanced Dissociation and Associated Surface Structure of Anesthetic Propofol at the Water Interface: Vibrational Sum Frequency Generation Study. *Phys Chem Chem Phys* **2021**, *23*, 24646–24651.

(18) Shamay, E. S.; Buch, V.; Parrinello, M.; Richmond, G. L. At the Water's Edge: Nitric Acid As a Weak Acid. *J Am Chem Soc* **2007**, *129*, 12910–12911.

(19) Wang, S.; Bianco, R.; Hynes, J. T. Depth-Dependent Dissociation of Nitric Acid at an Aqueous Surface: Car-Parrinello Molecular Dynamics. *J Phys Chem A* **2009**, *113*, 1295–1307.

(20) Baer, M. D.; Tobias, D. J.; Mundy, C. J. Investigation of Interfacial and Bulk Dissociation of HBr, HCl, and HNO3 Using Density Functional Theory-Based Molecular Dynamics Simulations. *J Phys Chem C* **2014**, *118*, 29412–29420.

(21) Galib, M.; Hanna, G. Molecular Dynamics Simulations Predict an Accelerated Dissociation of H2CO3 at the Air-Water Interface. *Phys Chem Chem Phys* **2014**, *16*, 25573–25582.

(22) Murdachaew, G.; Nathanson, G. M.; Gerber, R. B.; Halonen, L. Deprotonation of Formic Acid in Collisions with a Liquid Water Surface Studied by Molecular Dynamics and Metadynamics Simulations. *Phys Chem Chem Phys* **2016**, *18*, 29756–29770.

(23) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-Kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics. *Comput Phys Commun* **2018**, *228*, 178–184.

(24) Zhang, Y.; Wang, H.; Chen, W.; Zeng, J.; Zhang, L.; Wang, H.; E, W. DP-GEN:

A Concurrent Learning Platform for the Generation of Reliable Deep Learning Based Potential Energy Models. *Comput Phys Commun* **2020**, 107206.

(25) Yang, M.; Bonati, L.; Polino, D.; Parrinello, M. Using Metadynamics to Build Neural Network Potentials for Reactive Events: The Case of Urea Decomposition in Water. *Catalysis Today* **2021**, *387*, 143–149.

(26) Galib, M.; Limmer, D. T. Reactive Uptake of N2O5 by Atmospheric Aerosol Is Dominated by Interfacial Processes. *Science* **2021**, *371*, 921–925.

(27) Zeng, J.; Cao, L.; Xu, M.; Zhu, T.; Zhang, J. Complex Reaction Processes in Combustion Unraveled by Neural Network-Based Molecular Dynamics Simulation. *Nat Commun* **2020**, *11*, 5713.

(28) Seinfeld, J. H.; Pandis, S. N. *Atmospheric Chemistry and Physics*; John Wiley & Sons, 2016; p 1152.

(29) Rumble, J., Ed. *CRC Handbook of Chemistry and Physics*; CRC Press, 2021; Vol. 102nd Ed.

(30) Martins-Costa, M. T. C.; Anglada, J. M.; Francisco, J. S.; Ruiz-Lopez, M. F. Reactivity of Atmospherically Relevant Small Radicals at the Air-Water Interface. *Angew. Chem. Int. Ed.* **2012**, *124*, 5509–5513.

(31) Das, S.; Imoto, S.; Sun, S.; Nagata, Y.; Backus, E.; Bonn, M. Nature of Excess Hydrated Proton at the Water-Air Interface. *J Am Chem Soc* **2020**, *142*, 945–952.

(32) Yu, C.-C.; Imoto, S.; Seki, T.; Chiang, K.-Y.; Sun, S.; Bonn, M.; Nagata, Y. Accurate Molecular Orientation at Interfaces Determined by Multimode Polarization-Dependent Heterodyne-Detected Sum-Frequency Generation Spectroscopy Via Multidimensional Orientational Distribution Function. *J Chem Phys* **2022**, *156*, 094703.

(33) Minofar, B.; Vacha, R.; Wahab, A.; Mahiuddin, S.; Kunz, W.; Jungwirth, P. Propensity for the Air/Water Interface and Ion Pairing in Magnesium Acetate Vs Magnesium Nitrate Solutions: Molecular Dynamics Simulations and Surface Tension Measurements. *J Phys Chem B* **2006**, *110*, 15939–15944.

(34) Thomas, J.; Roeselová, M.; Dang, L.; Tobias, D. Molecular Dynamics Simulations of the Solution-Air Interface of Aqueous Sodium Nitrate. *J Phys Chem A* **2007**, *111*, 3091–3098.

(35) Otten, D. E.; Petersen, P. B.; Saykally, R. J. Observation of Nitrate Ions at the Air/water Interface by UV-Second Harmonic Generation. *Chem Phys Lett* **2007**, *449*, 261–265.

(36) Brown, M. A.; Winter, B.; Faubel, M.; Hemminger, J. C. Spatial Distribution of Nitrate and Nitrite Anions at the Liquid/Vapor Interface of Aqueous Solutions. *J Am Chem Soc* **2009**, *131*, 8354–8355.

(37) Tabe, Y.; Kikkawa, N.; Takahashi, H.; Morita, A. Surface Acidity of Water Probed by Free Energy Calculation for Trimethylamine Protonation. *J Phys Chem C* **2014**, *118*, 977–988.

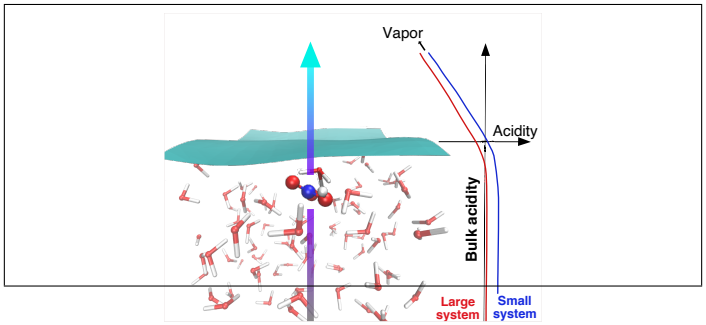
(38) Luo, M.; Wauer, N.; Angle, K.; Dommer, A.; Song, M.; Nowak, C.; Amaro, R.; Grassian, V. Insights into the Behavior of Nonanoic Acid and Its Conjugate Base at the Air/water Interface Through a Combined Experimental and Theoretical Approach. *Chem Sci* **2020**, *11*, 10647–10656.

(39) Enami, S.; Hoffmann, M. R.; Colussi, A. J. Proton Availability at the Air/Water Interface. *J. Phys. Chem. Lett.* **2010**, *1*, 1599–1604.

(40) Anglada, J.; Martins-Costa, M.; Francisco, J.; Ruiz-López, M. Reactivity of Undissociated Molecular Nitric Acid at the Air-Water Interface. *J Am Chem Soc* **2021**, *143*, 453–462.

(41) Rovelli, G.; Jacobs, M. I.; Willis, M. D.; Rapf, R. J.; Prophet, A. M.; Wilson, K. R. A Critical Analysis of Electrospray Techniques for the Determination of Accelerated Rates and Mechanisms of Chemical Reactions in Droplets. *Chem Sci* **2020**, *11*, 13026–13043.

Graphical TOC Entry



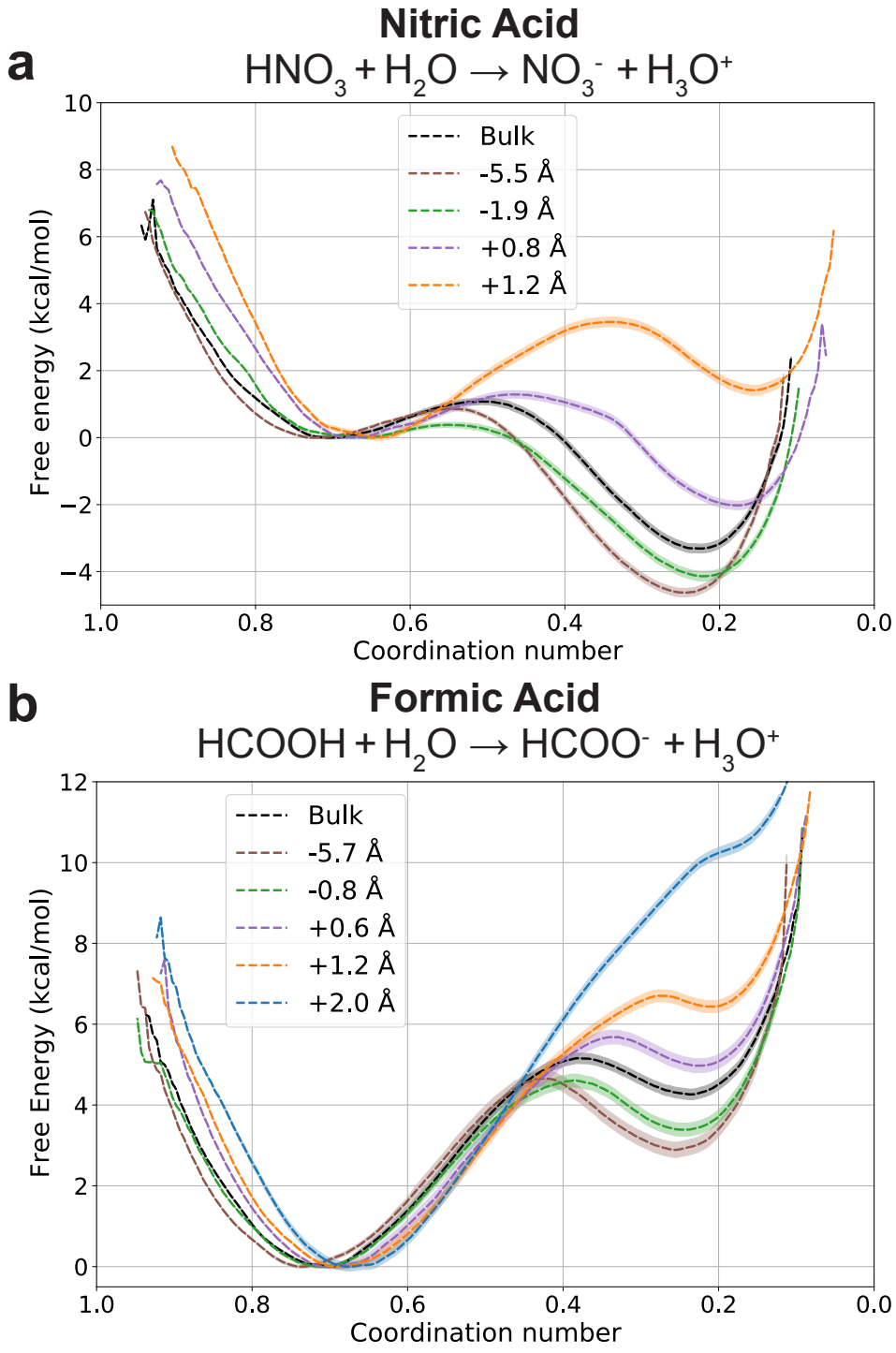


Figure 1: Dissociation free energy profiles (dashes) along the hydrogen coordination number around the acid oxygen atoms for a) nitric acid and b) formic acid, in the bulk and at a series of average distances between the nitric acid nitrogen atom (resp. the formic acid carbon atom) and the instantaneous air-water interface (see SI, distances taken to be positive on the vapor side and negative on the liquid side). Thick colored lines represent the 95% confidence interval.

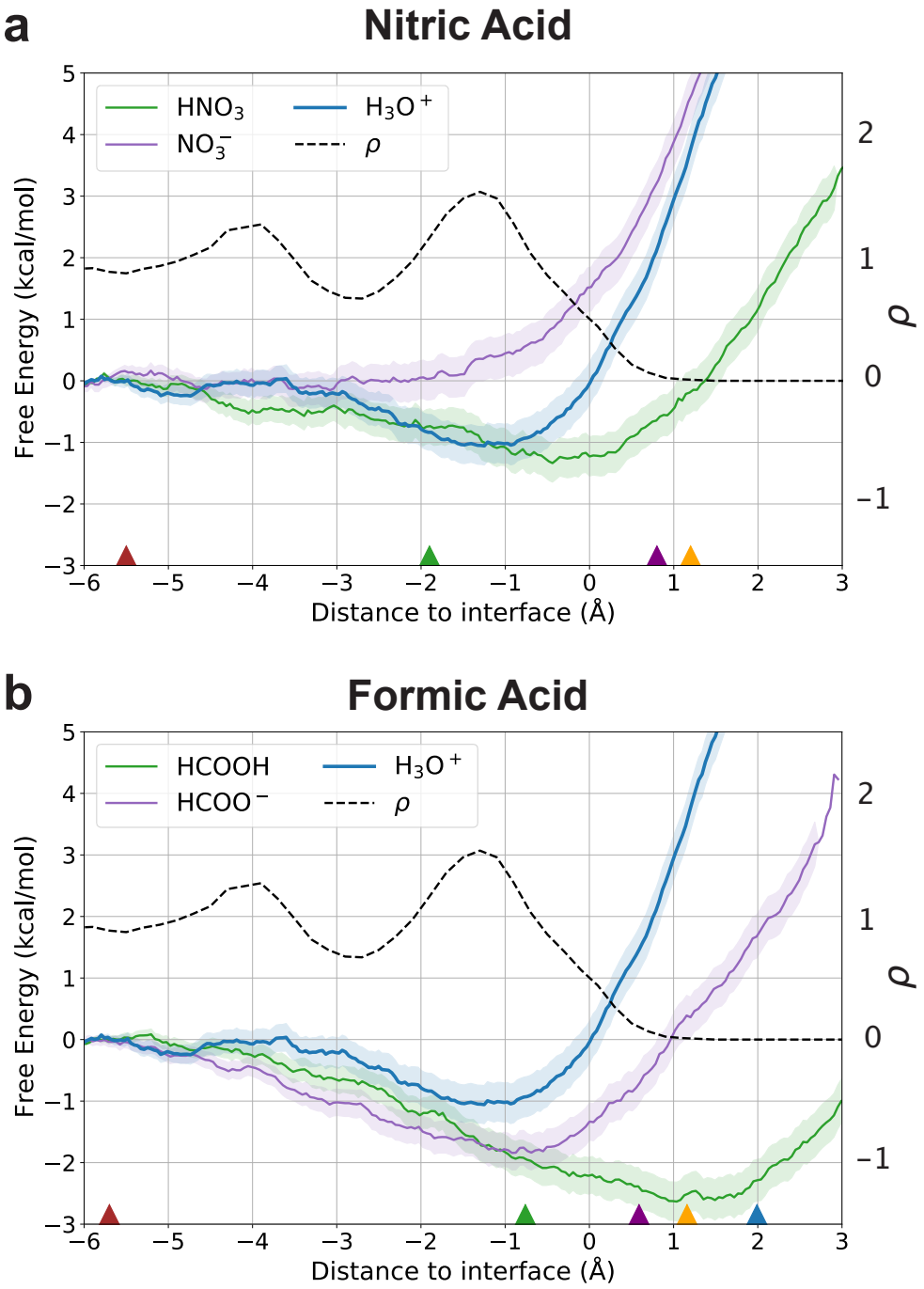


Figure 2: Hydration free energy profiles along the average distance to the instantaneous air-water interface a) for nitric acid, nitrate ion and hydronium ion, and b) for formic acid, formate ion and hydronium ion. The water density profile (dashes) shows the molecular layers. The triangles along the horizontal axis indicate the fixed positions used for the dissociation profiles in Fig. 1.

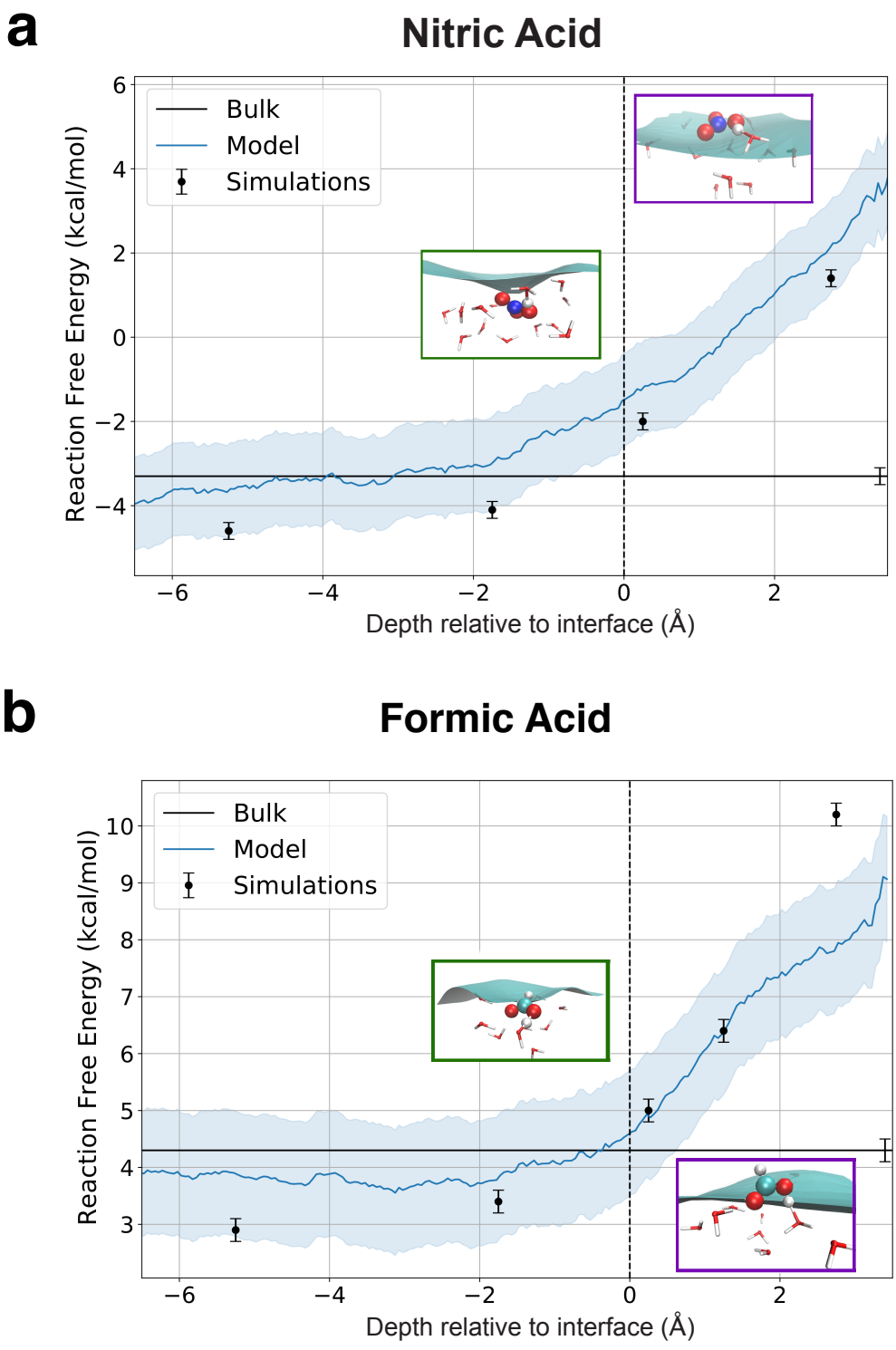


Figure 3: Acid dissociation free energies along the depth d relative to the average interface (vertical dashes), from the dissociation free energy profiles at fixed acid locations shown in Fig. 1 (black dots) and from the local solvation model eq. 1 (blue line) using the solvation free energies in Fig. 2 and the simulation slab thickness L=15.5 Å. Bulk references are shown by the horizontal lines. Uncertainties represent the 95% confidence interval. Insets show typical molecular configurations below (green, d=-1.75 Å) and above (purple, d=+0.25 Å) the interface.

Supplementary Information

Acids at the edge: why nitric and formic acid dissociations at air-water interfaces depend on depth and on interface specific area

Miguel de la Puente, Rolf David, Axel Gomez, and Damien Laage*

PASTEUR, Department of Chemistry, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France

E-mail: damien.laage@ens.psl.eu

Contents

Molecular Dynamics Simulations

All molecular dynamics simulations were performed in the canonical (NVT) ensemble with the LAMMPS code.[1] The 300 K target temperature was maintained with a Nose-Hoover thermostat[2–4] with a time constant of 0.1 ps. All trajectories were propagated with a 0.25 fs timestep. Two systems were considered: a bulk-like system comprising one solute (nitric or formic acid) with 128 water molecules in a 3D-periodic box of 16x16x16 $Å^3$, and an interface-like system consisting of the same molecules in a slab geometry placed at the center of a 16x16x48 $Å^3$ 3D-periodic box, with an approximately 32 Å-thick vacuum layer.

Deep Neural Network Potentials

Neural Network Potential Architecture and Training Parameters

Our DeePMD-kit[5] neural network potentials (NNP) consist of two deep neural networks. The first, *embedding* network builds invariant atom-centered descriptors from atomic configurations and is made of three layers of 20, 40 and 80 neurons. The second, *fitting* network whose input is the result of the embedding network yields the atomic energies and forces in a given configuration and is made of four 240 neuron layers. The embedding is performed according to the *smooth edition* scheme[6] with conservation of radial and angular information within a 6 Å cut-off. A cosine-type weighting scalar function is applied beyond 5.5 Å to ensure a smooth cut-off. While long-range electrostatic effects could be caused by the asymmetric interfacial environment studied in this work, short-ranged descriptions have been shown[7] to adequately reproduce the water liquid-vapor equilibrium, and we further mitigated this approximation by including interfacial configurations in the training set.

All neural network parameters are determined with a back-propagation training procedure.[5] The neural network is trained by minimizing a loss function applied to the root mean squared errors on the atomic energies and forces predicted by the network (compared to the reference calculation on the same configurations). This minimization is performed with an exponential learning rate that decays from 1×10^{-3} at the beginning of the procedure to a final value of 1×10^{-8}. Adjustable prefactors to the energy and force error terms are used, which vary over the training from 0.01 to 0.1 and from 1000 to 1 respectively. The total number of training batches (minimization steps) ranges from 4×10^5 for the NNP of the initial iterations (see next subsection) to 1.1×10^6 for the production models. Batches of 1 and 4 configurations were used respectively for training and on-the-fly testing during the loss function minimization.

Iterative Construction of the Training Set

NNP are trained over a training set which consists in an ensemble of atomic configurations ("inputs") for which the atomic energies and forces ("labels") are calculated at a reference level of theory, here BLYP-D3/TZV2P (see description of labelling below). Here, such a set of configurations was first generated by propagating trajectories with a different level of theory. Our initial training set was constructed with configurations extracted from trajectories propagated with DFTB2/MIO-1.1.[8,9] This choice was motivated by the reduced computational cost of this method and its ability to describe chemical reactions. However, the DFTB2 equilibrium distribution of atomic configurations is fairly distinct from that of our BLYP-D3/TZV2P reference level. A second training set was therefore obtained by propagating new trajectories with this initial NNP to obtain configurations closer to the reference equilibrium distribution (the initial DFTB2 configurations were not included in this second set). This second training set was then used to train a NNP and to start a concurrent learning procedure inspired by the DP-GEN[10] scheme.

This iterative procedure leads to a training set representative of the system's phase space region of interest, which is essential for a correct description by the NNP. The iterations consisted of three phases, exploration, labelling and training, that we now describe.

In the training phase, 3 NNPs were trained on the current iteration's training set with the procedure described in the previous subsection. They used the same training set but a different random initialization of their parameters.

The exploration phase consisted of one or several (up to 15) molecular dynamics trajectories. Forces were propagated using the LAMMPS software[1] interfaced with the DeePMD plugin. Biasing constraints were used on some iterations with the PLUMED[11] package. The length of the exploration trajectories ranged from 10 ps for initial iterations to 150 ps for the final ones. At the end of every exploration phase, candidate configurations to be included in the training set were selected. Configurations were considered as candidates when the maximum deviations in the atomic forces predicted by the three NNPs were in the $[0.1, 0.5]$

eV/Å range (such that the configurations are physically meaningful while not satisfactorily described by the NNP). A maximum number of 500 candidate configurations per trajectory were selected randomly.

In the labelling phase, we computed the energies and atomic forces of the selected candidate configurations with the BLYP[12,13] functional with D3 dispersion corrections,[14] and using GTH pseudo-potentials[15,16] for core electrons and a triple-ζ basis set (TZV2P[17]) for valence electrons. All electronic structure calculations were performed with the CP2K[18] software. These inputs (atomic configurations) and labels (energies and forces) were then added to the total training set.

We considered that the NNP were converged with respect to training set size when no candidate configurations were found at the end of an exploration phase. We first considered convergence of the NNP only for bulk configurations, which was reached after 6 and 9 iterations for nitric and formic acids respectively. We then started adding interface configurations, and convergence was reached for 5 (nitric) and 6 (formic) additional iterations. The final training sets consisted of 8340 configurations for nitric acid and 11608 configurations for formic acid.

The learning curves in Fig. S1a-b show that the Root-Mean-Square Error (RMSE) on atomic forces relative to the standard deviation of the force distribution reaches a plateau for both the training set and for an independent test set not used for training. In addition, the final NNP atomic forces on the test set exhibit a very strong correlation with reference DFT calculations (Fig. S1c-d). These results show that the NNP has properly converged and yields results of BLYP-D3/TZV2P quality.

Neural Network Potentials Performances

NNP-propagated molecular dynamics simulations were approximately 2,000 times faster than DFT-based simulations for the machines that we used. The production rate was $\simeq$ 2 ns/day for NNP-MD on 1 Nvidia Tesla V100 SXM2 GPU with 16 GB of RAM compared to $\simeq$ 1

ps/day of DFT-MD at the BLYP-D3/DZVP level on a node with 2 AMD EPYC 7542 CPUS (for a total of 64 CPU cores), both with a timestep of 0.25 fs. This allowed us to perform long reactive molecular dynamics simulations of DFT quality, with an accrued total simulation time around 60 ns, which would not have been possible using traditional DFT-MD.

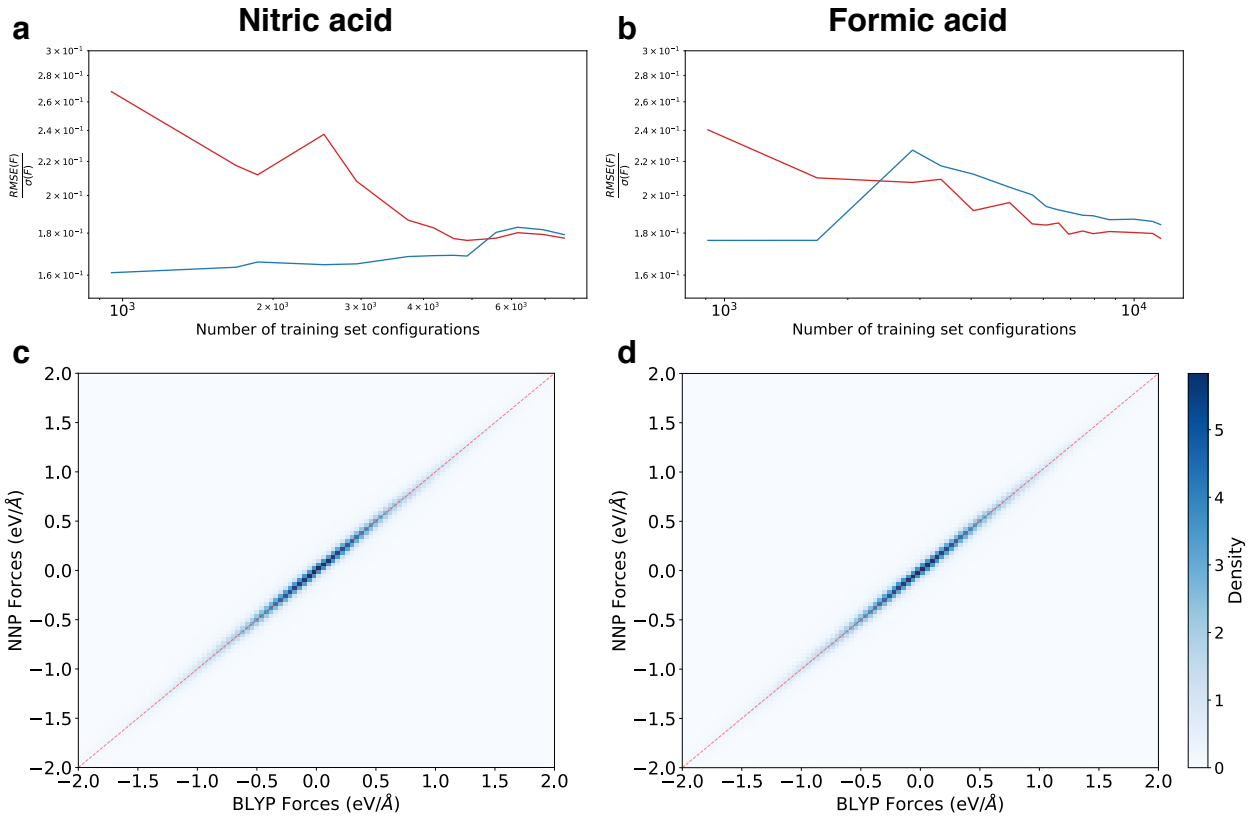


Figure S1: a) Nitric acid NNP concurrent learning curves showing the RMSE of atomic forces relative to their standard deviation as the training set size is increased, calculated respectively over the entire training set (blue) and on a test set (red) consisting in 600 bulk and 600 interface configurations, equally distributed between reactant, product and transition state configurations. b) Same as a) for formic acid. c) Density map of correlations between atomic forces from reference BLYP-D3 calculations and NNP for nitric acid calculated on the 1,200 bulk and interface configuration test set (mean average error of 0.065 eV/Å) d) Same as c) for formic acid (mean average error of 0.066 eV/Å).

Free Energy Profile Calculations

All free energy calculations were performed with Umbrella Sampling. The global free energy profiles were obtained from the series of window simulations with a homemade weighted histogram analysis method (WHAM)[19] implementation. Uncertainties on the free energies were calculated as described by Zhu and Hummer.[20]

Reaction Free Energy Profiles

The deprotonation reactions were described with a collective variable which measures the coordination number n of all labile hydrogen atoms in the system around the acid oxygen atoms:

$$n = \sum_{i \in \{O_{acidic}\}} \sum_{j \in \{H_{labile}\}} \frac{1 - \left(\frac{r_{ij}}{r_0}\right)^6}{1 - \left(\frac{r_{ij}}{r_0}\right)^{12}} \tag{S1}$$

The cut-off radius r_0 was set to 1.1 Å so that covalent interactions were included with a minimal perturbation coming from the hydrogen bonds around the solute. With this cut-off value, the reactant (protonated state) coordination is $n \simeq 0.8$ and the product (deprotonated state) coordination is $n \simeq 0.2$, consistent with the deprotonation.

The Umbrella Sampling calculations were performed with a total of 40 separate windows: a first series of 16 windows used a 0.05 coordination increment spanning the $[0.1, 0.9]$ interval, with 4 additional windows around the expected transition state position (0.375, 0.425, 0.475 and 0.525), and 20 supplementary windows centered around coordination values corresponding to the middle of pairs of adjacent windows (therefore shifted by 0.025 or 0.0125 with respect to the 20 original ones). Coordination was constrained with harmonic potentials of 15 eV per squared coordination unit. For deprotonation reactions at the air-water interface, an additional harmonic constraint of 9.72 eV/Å^2 was added on the distance between the centers of mass of the solute and of the water slab (this force constant value was successfully used in other prior work[21]). For the bulk reaction, 40 ps-long simulations were used for every target value of the collective variable. At the interface, 256 ps-long simulations were used

when the target value of the collective variable was inferior to 0.3 (in order to ensure that the free hydronium ion can explore the entire slab), and 50 ps-long trajectories were used when the target coordination value was above 0.3.

Solvation Free Energy Profiles

Solvation free energy profiles were first calculated along the distance d_{com} between the center of mass of the water slab (which was fixed by removing the total linear momentum of the system every 100 simulation steps) and the central atom of each solute (N atom for nitric acid and nitrate ion, C atom for formic acid and formate ion, and O atom for the hydronium ion). Harmonic constraints of 0.5 eV/Å^2 were applied to bias this distance in the Umbrella Sampling calculations. The free energy profiles were calculated with a total of 80 separate windows with target distances separated by 0.125 Å and covering the $]0.0, 10.0]$ Å interval. In the case of formic acid and formate, 20 extra windows were added to assess the free energy convergence: their centers were shifted by 0.0625 Å with respect to the existing windows and spanned the $]0.0, 2.5]$ Å interval, but no change in solvation free energies was observed and the cumulated errors were reduced by less than 0.1 kcal/mol. For the acidic and basic forms of the two solutes, harmonic walls of 20 eV force constant were added on the upper and lower values of the coordination collective variable respectively, to prevent any chemical reaction. For the hydronium ion solvation free energy, in order to prevent any proton transfer, every window was started from an *Eigen* configuration (with the excess proton localized on one oxygen atom) and we applied a harmonic wall with a 20 eV force constant on the coordination number around the central oxygen atom due to the three hydrogen atoms of the initial hydronium (analogous to n in eq S1). Trajectories were propagated for 50 ps in each window.

While the distance to the center of mass of the water slab (d_{com}) allowed efficient on-the-fly biasing during the umbrella sampling calculations, representing the free energy profiles along the distance from each solute to the interface provides a more insightful description.

We have considered both the distance to the average Gibbs dividing surface (GDS), defined as the plane at which the average water density reaches half its bulk value, and the distance to the instantaneous air-water interface,[22] defined as a constant density surface in a single configuration.

For the calculation of solvation free energy profiles for each species along the distance to the interface, the identification of molecular layers, and the selection of solutes lying in the topmost layer to study their orientation at the interface (see below), we have considered the average distance to the instantaneous interface. While time fluctuations are averaged, it captures the local spatial distortion of the interface due to the presence of the solute. We calculated the instantaneous air-water interface (see procedure below) at each step of the umbrella sampling window trajectories used for the solvation free energy profile calculations, and we determined the correlation between the distance d to the GDS and the distance d_{inst} between the central atom of each solute (N atom for nitric acid and nitrate ion, C atom for formic acid and formate ion, and O atom for the hydronium ion) and the instantaneous interface. As shown in Fig. S2, these correlations can be approximated by affine functions. We therefore determined approximate relations between d and d_{inst} (see insets in Fig. S2) that we then used to obtain the hydration free energy profiles presented in Fig. 2.

For the dissociation free energy profile along the distance to the interface in Fig. 3, since several chemical species are involved which induce different local distortions of the interface, we considered the distance to the Gibbs dividing surface. The GDS provides a local frame of reference common to the different chemical species under consideration. The solvation free energy profiles along the depth d of a solute with respect to the GDS can be straightforwardly obtained from the profiles along the distance to the slab center of mass by a simple translation of the local frame of reference (see Fig. S8). For our two systems, the GDS were found to lie respectively 6.75 Å above and below the center of mass of the slab, so that $d = d_{com} - 6.75$ Å.

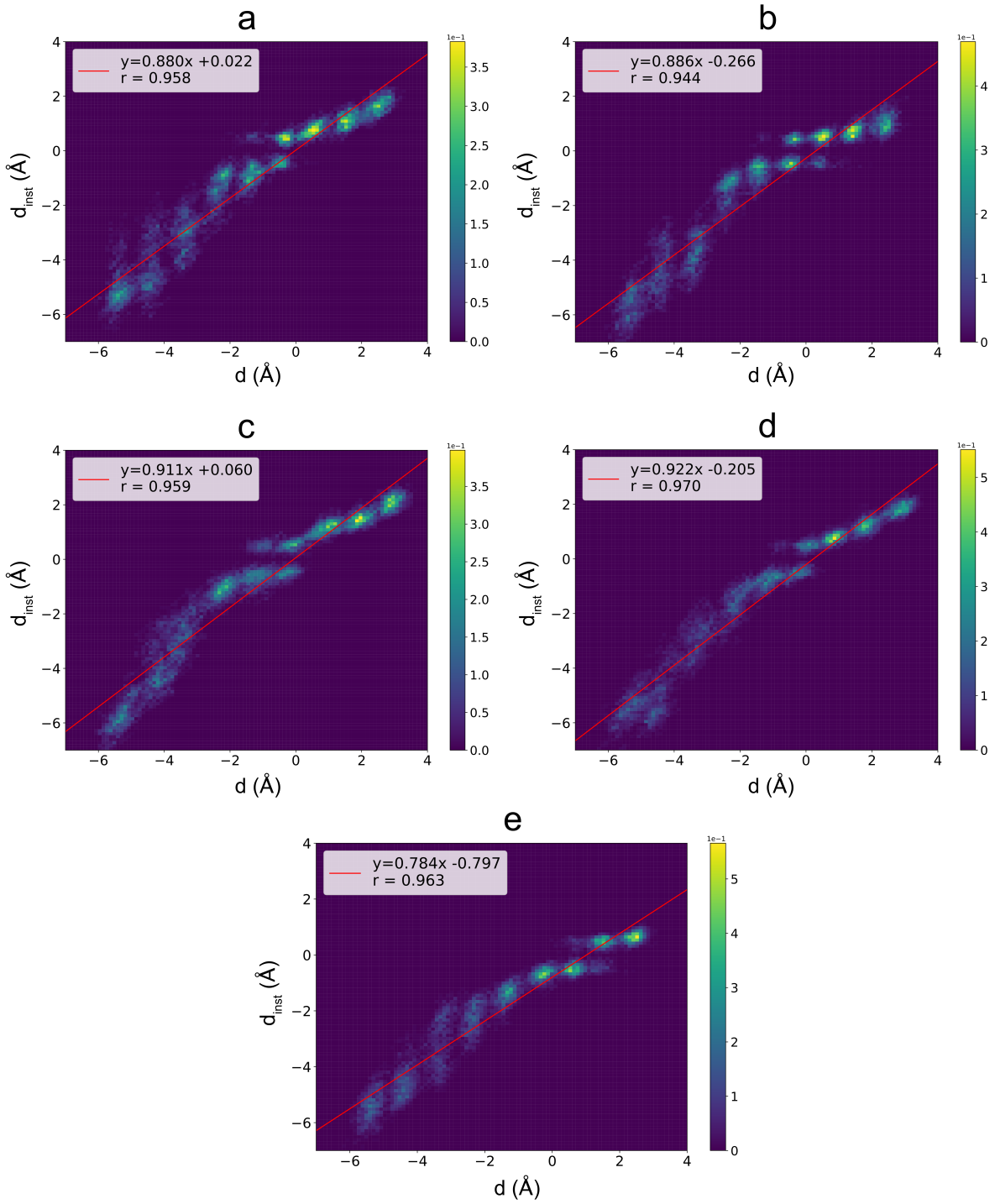


Figure S2: Correlations between the depth d relative to the average interface and the distance to the instantaneous interface d_{inst} for nitric acid (a), nitrate (b), formic acid (c), formate (d) and hydronium ion (e).

Instantaneous Air-Water Interface

The instantaneous air-water interface was calculated following the procedure described in ref. 22. This interface was defined as the height (for fixed x and y coordinates in the simulation box frame) at which the continuous water density (defined as the sum of normalized Gaussian potentials of 2.4 Å standard deviation centered on the water oxygen atoms) reaches a value of 0.016 $Å^{-3}$, half of its bulk value for water at room temperature and ambient pressure. We computed the interface with horizontal grids of 16x16 nodes. The distance of a solute to this interface is approximated as the distance from the central atom of the solute (see Fig. S2) to the closest grid node.

Dissociation Free Energy at Different Depths

We report the dissociation reaction free energy $\Delta_r G(d)$ (eq. 1) values for nitric (Table S1) and formic (Table S2) acids at different depths in the water slab. We also report the average distance between the transition state of each system and the instantaneous interface ($d_{inst}(TS)$), which informs us about solvation at a given depth in the slab. The bulk pK_A is determined by the reaction free energy $\Delta_r G^\circ$ in the standard state with solute concentrations of 1 mol/L. We obtain the latter from the reaction free energy $\Delta_r G$ calculated in our system at a concentration of 1 acid molecule in 128 water molecules as

$$pK_A = 1/\left[\ln(10) k_B T\right] \Delta_r G^\circ = 1/\left[\ln(10) k_B T\right] \left[\Delta_r G - k_B T \ln\left(C/C^\circ\right)\right] \tag{S2}$$

with $C \simeq 55/128$ mol/L and $C^\circ = 1$ mol/L.

Table S1: Local reaction free energies of nitric acid at different relative depths in the water slab. The width of the 95% confidence interval on the free energy values is 0.2 kcal/mol. Positive (negative) values of $d_{inst}(TS)$ refer to solute center of mass positions above (below) the instantaneous interface.

d [Å]	Bulk	−5.25	−1.75	+0.25	+2.25
$d_{inst}(TS)$ [Å]	n.a.	−5.5	−1.9	+0.8	+1.2
$\Delta_r G(d)$ [kcal/mol]	−3.3	−4.6	−4.1	−2.0	+1.4

Table S2: Local reaction free energy of formic acid at different relative depths in the water slab. The width of the 95% confidence interval on the free energy values is 0.2 kcal/mol. Positive (negative) values of $d_{inst}(TS)$ refer to solute center of mass positions above (below) the instantaneous interface.

d [Å]	Bulk	-5.25	-1.75	$+0.25$	$+1.25$	$+2.25$
$d_{inst}(TS)$ [Å]	n.a.	-5.7	-0.8	$+0.6$	$+1.2$	$+2.0$
$\Delta_r G(d)$ [kcal/mol]	4.3	2.9	3.4	5.0	6.4	10.2

System Size Effects

To examine possible size effects, we have repeated our molecular dynamics simulations for a system that is twice larger and contains 256 water molecules.

We have first calculated the bulk dissociation free energy profile for one acid molecule solvated in a (19.8 Å)^3 box of 256 water molecules, following the same procedure as described for the system with 128 water molecules (same reaction coordinate, umbrella sampling procedure and simulation lengths). The obtained bulk dissociation free energies are -3.8 kcal/mol for nitric acid and +3.2 kcal/mol for formic acid.

We have also repeated our calculations for slab systems with 256 water molecules, which have the same interface area as the slabs with 128 water molecules and are twice thicker, in a simulation box of 16×16×96 Å. Dissociation free energy profiles at selected depths relative to the interface were obtained with the same procedure as described for the slabs with 128 water molecules (depths of -1.70 Å and +2.45 Å were used for nitric acid, and of +0.55 Å and +2.25 Å for formic acid).

The dissociation free energy predicted by our solvation model eq. 1 was obtained from the solvation free energy profiles determined from the 128 water molecule simulations (Fig. 2) for distances to the interface smaller than 6.75 Å which we extended by zero padding for distances greater than 6.75 Å(i.e. assuming that the bulk-like solvation plateau is reached). The slab thickness was L_{256} =31.0 Å. The dissociation free energy predicted by the model was then obtained as

$$\Delta_r G^{256}(d) = \Delta_r G^{256}_{\text{bulk}} + \Delta G^{B,128}_{\text{hyd}}(d) + \langle \Delta G^{H,128}_{\text{hyd}} \rangle_{L_{256}} - \Delta G^{A,128}_{\text{hyd}}(d) - \langle \Delta G^{W,128}_{\text{hyd}} \rangle_{L_{256}} \quad \text{(S3)}$$

The comparison of the dissociation free energies for nitric acid and formic acid from the model and from the direct calculation in Fig. S3 shows that the agreement remains excellent in this larger system, which suggests that our model can be used to describe local acidity in a broad range of system sizes.

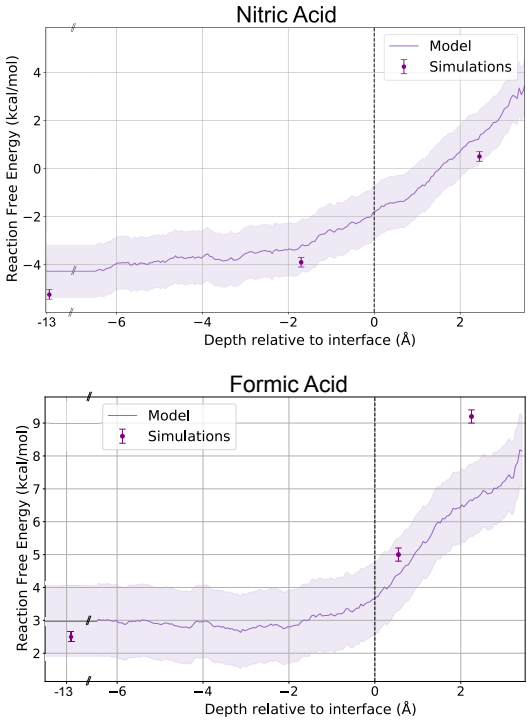


Figure S3: Acid dissociation free energies along the depth d relative to the average interface (vertical dashes) for one acid and 256 water molecules, from the direct calculation dissociation free energy profiles at fixed acid locations (purple dots) and from the local solvation model eq. 1 (purple line) using the solvation free energies determined from our calculations on systems with 128 water molecules in Fig. 2 and the 256 molecule slab thickness L=31 Å. Uncertainties represent the 95% confidence interval.

Solute Orientation at the Air-Water Interface

We determined the orientations of the protonated reactant (R), deprotonated product (P) and transition state (TS) forms of the solutes at different relative depths relative to the interface. These distributions are shown in Fig. S4 and S5, and some typical configurations are presented in Fig. S6.

Regarding nitric acid, when below and immediately above the interface ($d = -1.75$Å and $+0.25$ Å), the molecular plane lies parallel to the interface, so as to maximize the number of H-bonds with the water slab. Deprotonation occurs in this configuration with a water molecule in the first molecular layer acting as the proton acceptor. When nitric acid is well above the interface ($d = +2.25$ Å), its orientational preference is less pronounced, parallel and

perpendicular configurations with respect to the interface are accessible, and deprotonation occurs to a water molecule that leaves the slab due to the strong H-bond with the acidic nitric acid OH group.

As for formic acid, it is always oriented in order to minimize the interactions between its CH moiety and the water slab, which is consistent with the expected amphiphilic behavior of this acid. A slightly tilted configuration is found when the acid is beneath the interface ($d = -1.75$ Å) where the proton accepting water molecule lies in the first layer. Well above the interface ($d = +1.25$ Å), formic acid behaves like nitric acid: it can adopt different orientations and forms an H-bond with an interfacial water molecule to which it will transfer its proton above the interface.

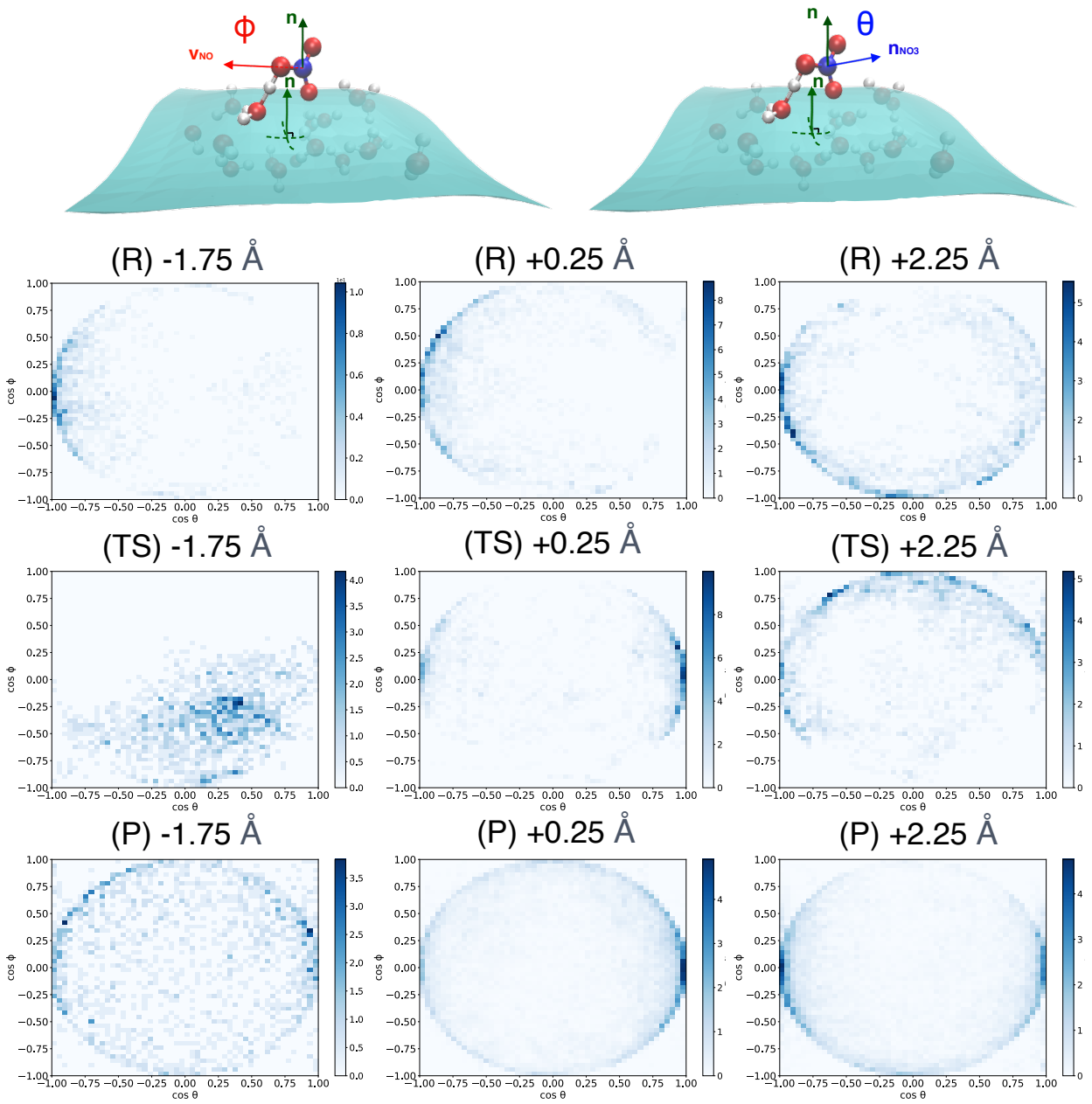


Figure S4: Normalized orientation probability density distributions of nitric acid in the reactant (R), transition state (TS) and product (P) forms at different relative depths from the air-water interface. The molecular orientation is defined by the ϕ angle between the NO bond unit vector v_{NO} (where the O atom is that which carries the hydrogen atom) and the vector n normal to the instantaneous interface at the location closest to the solute, and by the θ angle between n and the n_{NO3} unit vector normal to the plane defined by the three NO_3 oxygen atoms.

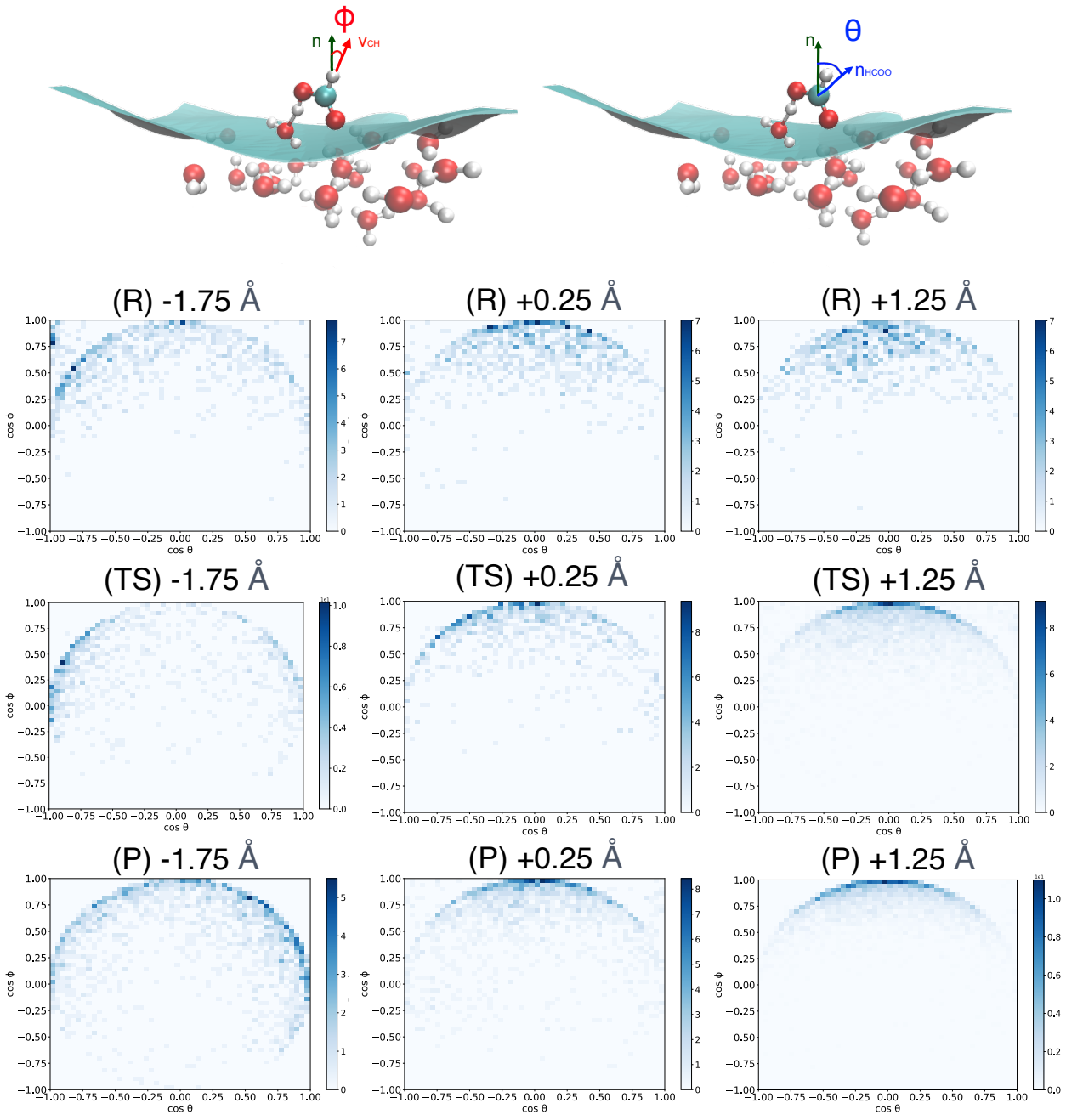


Figure S5: Normalized orientation probability density distributions of formic acid in the reactant (R), transition state (TS) and product (P) forms at different relative depths from the air-water interface. The molecular orientation is defined by the ϕ angle between the CH bond unit vector v_{CH} and the vector n normal to the instantaneous interface at the location closest to the solute, and by the θ angle between n and the unit vector n_{HCOO} normal to the plane defined by the COO atoms.

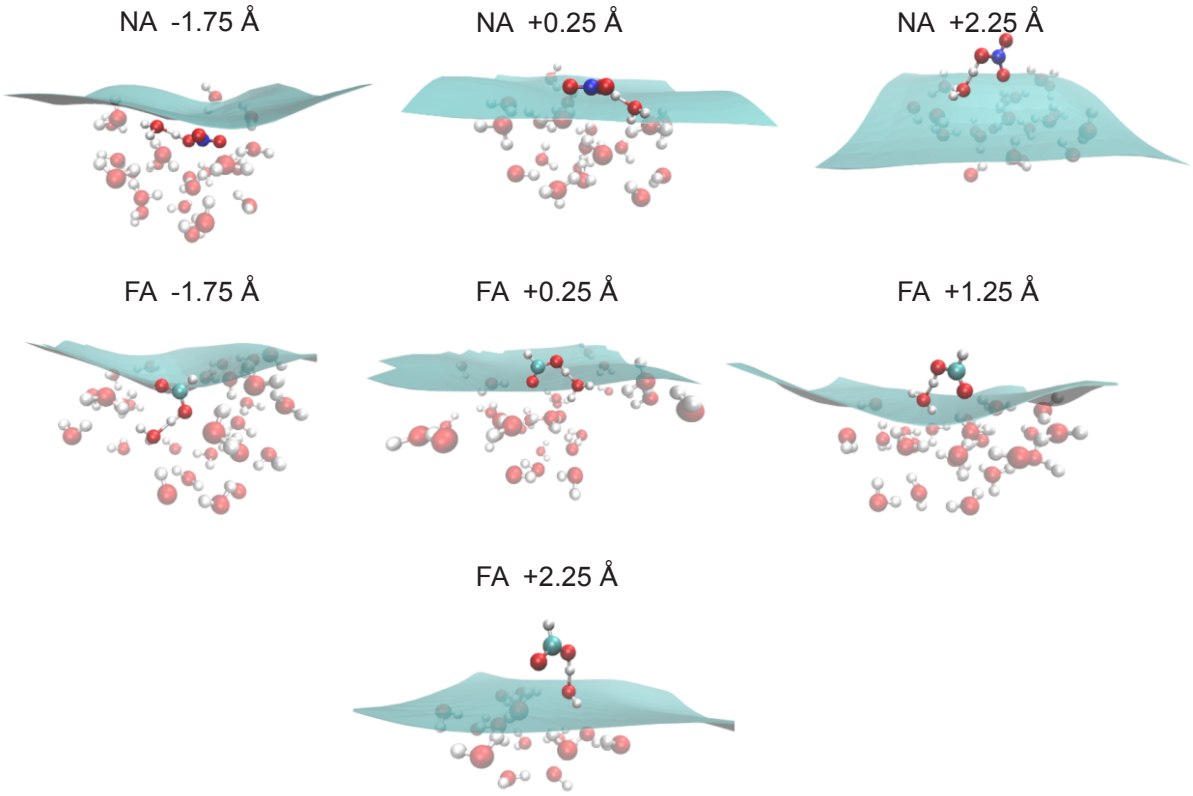


Figure S6: Typical configurations for nitric (NA) and formic (FA) acids at different relative depths from the air-water interface.

Shift in Formic Acid CH Stretch Vibrational Frequency upon Deprotonation in the Bulk and at the Air-Water Interface

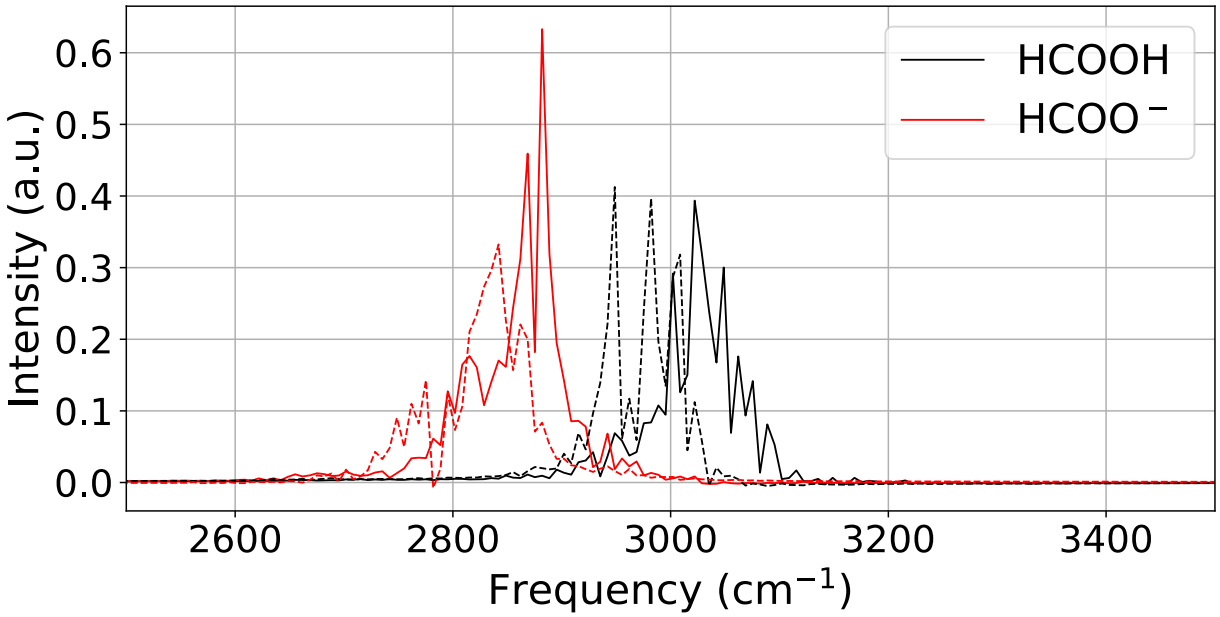


Figure S7: Calculated CH stretch vibrational density of states for formic acid and formate ion in the bulk (solid lines) and at the air-water interface (dashed lines).

Raman spectroscopy studies of formic acid and formate ion in bulk aqueous solution showed (see ref. 23 and references therein) that the CH stretch vibrational frequency exhibits a frequency red-shift of approximately 100 cm^{-1} upon deprotonation. This CH vibrational frequency has therefore been a key point in the analysis of SFG experiments,[23] to determine the protonation state of formic acid at the interface. However, it is typically assumed that the shift observed in the bulk is preserved at the interface. We have therefore determined the CH stretch vibrational density of states in formic acid and in the formate ion, both in the bulk and at the air-water interface. The vibrational densities of states were obtained from the Fourier transform of the CH bond length time correlation functions obtained from 50 ps-long NNP molecular dynamics trajectories propagated with a 0.25 fs timestep.

Our results in Fig. S7 first show that our NNP force field reproduces the CH stretch frequency red shift in the bulk almost quantitatively. In addition, it shows that at the interface, the CH stretch vibrational frequency exhibits a slight red shift with respect to the bulk (as could be expected, since the repulsion with surrounding water molecules that was present in the bulk disappears at the interface). However, this shift is almost an order of magnitude weaker than that associated with the chemical transformation, and is fairly similar for the acid and conjugated base forms. The frequency shift between formic acid and formate ion at the interface thus remains close to the bulk value, and this supports the use of the CH stretch frequency as a method to identify interfacial formic acid.

Local Acidity Model

System Definition

We consider the general acid-base reaction

$$\mathrm{AH} + \mathrm{H_2O} \rightleftharpoons \mathrm{A^-} + \mathrm{H_3O^+} \tag{S4}$$

taking place in a slab system, periodically replicated along the x and y directions, of thickness L and extending from $-L/2$ to $L/2$ along the z direction. In the following we note $p_{A,B}(z)$ the probability of finding the acid AH (A) and the conjugate base A^- (B) at position z, and $x^L_{A,B}$ the overall acid and base fractions in the slab.

At equilibrium, thermodynamics requires that chemical potentials be uniform in the slab, and defining local thermodynamic quantities can be ambiguous. Here we probe the local acidity of the acid AH by considering its degree of dissociation.

The acid and base fractions reported by a given experimental technique depend on its probing depth and on the weight given to each layer, and can be generally expressed as

$$x^{technique}_{A,B} = x^L_{A,B} \frac{\int_{-L/2}^{+L/2} dz \;\; p_{A,B}(z) w^{technique}(z)}{\int_{-L/2}^{+L/2} dz \;\; p_{A,B}(z)} \tag{S5}$$

where $w^{technique}(z)$ is the weighting function of the technique under consideration.

In the following, we will first determine the average acid and base fractions in the entire slab $x^L_{A,B}$ and then discuss the weighting functions corresponding to a series of experimental techniques, in order to compare how the reported acid and base fractions are affected by the system specific area and by the chosen technique.

Local Acidity in an Interfacial System

When the dissociation reaction occurs with the acid (A), base (B), hydronium ion (H) and reactant water molecule (W) all fixed at position z in the slab, the change in reaction free

energy with respect to its bulk value is due to the change in hydration free energy with respect to the bulk situation $\Delta G_{hyd}(z) = G(z) - G_{bulk}$ for the acid, base, hydronium ion and water molecule, that we assume to be independent,

$$\Delta_r G^{\text{fixedH}}(z) = \Delta_r G_{\text{bulk}} - \Delta G^A_{\text{hyd}}(z) - \Delta G^W_{\text{hyd}}(z) + \Delta G^B_{\text{hyd}}(z) + \Delta G^H_{\text{hyd}}(z) \tag{S6}$$

The more relevant situation is that where acid and base are at position z in the slab, but the reactant water molecule and hydronium ion product can freely diffuse throughout the slab. The reaction free energy averaged over the locations of both species is then

$$\Delta_r G(z) = \Delta_r G_{\text{bulk}} - \Delta G^A_{\text{hyd}}(z) - \langle \Delta G^W_{\text{hyd}} \rangle_L + \Delta G^B_{\text{hyd}}(z) + \langle \Delta G^H_{\text{hyd}} \rangle_L \tag{S7}$$

Equation 1 in the main text is then simply obtained by translation along z, using the relation $d = z - z_{GDS}$. The hydronium ion average hydration free energy change with respect to the bulk is

$$\left\langle \Delta G^H_{hyd} \right\rangle_L = -k_B T \ln \left[\frac{1}{\Lambda} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz \, e^{-\Delta G^H_{hyd}(z)/k_B T} \right] + k_B T \ln \left[\frac{1}{\Lambda} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz \right] \tag{S8}$$

with Λ the thermal wavelength. For water this can be expressed in terms of the density profile $\rho(z)$ as

$$\left\langle \Delta G^W_{hyd} \right\rangle_L = -k_B T \ln \left[\frac{1}{\Lambda} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz \, \frac{\rho(z)}{\rho_{bulk}} \right] + k_B T \ln \left[\frac{1}{\Lambda} \int_{-\frac{L}{2}}^{\frac{L}{2}} dz \right] \tag{S9}$$

In the limit of an infinitely thick slab, the average hydration free energy changes in eqs. S8-S9 vanish because the interfacial region where $\Delta G^H_{hyd}(z) \neq 0$ and $\rho(z) \neq \rho_{bulk}$ becomes negligible with respect to the bulk-like region in the middle of the slab.

We note that a prior theoretical study[24] performed on the dissociation of a charged acid at the interface of an infinitely thick slab had underlined the important role of acid and base

solvation in the dissociation reaction free energy. In the limit of a vanishing surface/volume ratio for the aqueous phase, the $\left\langle \Delta G_{hyd}^{H,W} \right\rangle_L$ terms are negligible and Equation 1 in the main text becomes identical to the result obtained in ref. 24.

To obtain the hydration free energy differences between the location z and the bulk, we approximate the bulk hydration free energy by the hydration free energy in the center of the slab, which yields

$$\Delta G_{hyd}^{A,B,H,W}(z) = G^{A,B,H,W}(z) - G_{bulk}^{A,B,H,W} \simeq G^{A,B,H,W}(z) - G^{A,B,H,W}(z=0) \tag{S10}$$

Figure S8 shows the hydration free energy profiles across the slab, shifted such that $\Delta G_{hyd}^{A,B,H}(z=0) = 0$ (the water density profile is shown in Fig. 2 of the main text). The approximation used in eq. S10 is supported by the fact that the hydration free energy profile of each species reaches a plateau in the middle of the slab and by the excellent quantitative agreement with the simulation $\Delta_r G(z)$ shown in Fig. 3 of the main text. We stress that in the middle of the slab solvation is assumed to be bulk-like, but acidity is not bulk-like because the released hydronium ion can diffuse to and stabilize at the interface.

To extrapolate our model to slabs thicker than the slab used in our molecular dynamics simulations, we consider that in the region in the middle of the slab the hydration free energy differences for distances from the interface larger than these considered in Fig. S8 are zero.

Average dissociation equilibrium constant $\langle K_A \rangle_L$

The average dissociation constant for the dissociation reaction eq. S4 is given by the Law of Mass Action

$$\langle K_A \rangle_L = \frac{x_B^L \; x_H^L}{x_A^L x_W^L} \tag{S11}$$

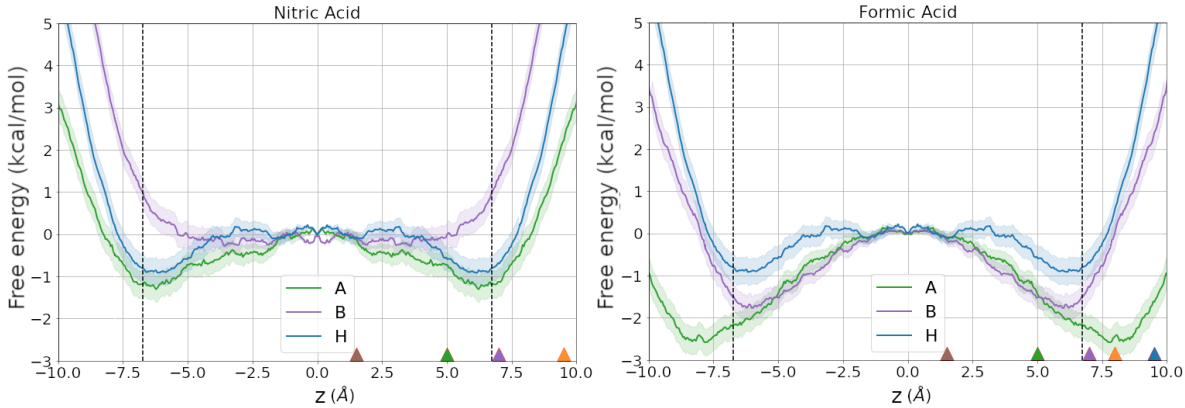


Figure S8: Hydration free energies $\Delta G_{hyd}^{A,B,H}(z)$ across the water slab for nitric (left) and formic (right) acid species, respectively for the acid (A), conjugate base (B) and hydronium ion (H). Colored triangles indicate the depths at which reaction free energies were explicitly calculated and reported in Fig. 1. The dashed black lines indicate the positions of the Gibbs dividing surfaces.

where the molar fraction for species i (among A, B, H and W) in the slab of thickness L is determined by the probability distribution integral across the slab,

$$x_i^L = \frac{1}{L}\int_{-L/2}^{L/2} \mathrm{d}z \;\; p_i(z) \tag{S12}$$

These probability distributions are given by the local free energy, which is the sum of the bulk free energy and of the hydration free energy change with respect to the bulk

$$p_i(z) = e^{-G^i(z)/k_BT} = e^{-G^i_{bulk}/k_BT} \;\; e^{-\Delta G^i_{hyd}(z)/k_BT} \tag{S13}$$

Combining eqs. S11–S13 yields

$$\langle K_A \rangle_L = e^{-\Delta_r G_{bulk}/k_BT} \frac{e^{-\left\langle \Delta G^H_{hyd}(z)\right\rangle_L/k_BT} e^{-\left\langle \Delta G^B_{hyd}(z)\right\rangle_L/k_BT}}{e^{-\left\langle \Delta G^A_{hyd}(z)\right\rangle_L/k_BT} e^{-\left\langle \Delta G^W_{hyd}(z)\right\rangle_L/k_BT}} \tag{S14}$$

which can be rewritten as

$$\langle K_A \rangle_L = K_A^{bulk} \;\; K_L \tag{S15}$$

with

$$K_L = \frac{e^{-\left\langle \Delta G^H_{hyd}(z) \right\rangle_L / k_B T} e^{-\left\langle \Delta G^B_{hyd}(z) \right\rangle_L / k_B T}}{e^{-\left\langle \Delta G^A_{hyd}(z) \right\rangle_L / k_B T} e^{-\left\langle \Delta G^W_{hyd}(z) \right\rangle_L / k_B T}} \tag{S16}$$

Equation S15 thus expresses the Law of Mass Action for the dissociation equilibrium in the slab.

Finally, using eq. S11 and the conservation of matter, the average acid molar fraction x_A^L in the slab is thus the solution of

$$\langle K_A \rangle_L = \frac{(1 - x_A^L)^2}{x_A^L} \tag{S17}$$

Dependence of measured acid/base ratios on system specific area and experimental probing depth

Experimental probing depth. Experimental techniques that have been used to probe the acid/base ratios at air-water interfaces have different probing depths. We now define simplified weighting functions to be used in eq. S5 to model these probing depths. Our goal is to stress the impact of different probing depths but we do not aim to provide a quantitative description of the experiments, which would require a specific calculation of the actual signal collected by each technique.

We first consider vibrational Sum Frequency Generation (SFG) spectroscopy. SFG is surface specific and probes the acid/base ratio within a very thin layer at the air-water interface l, on the order of a molecular layer. If one assumes for simplicity that the signal is uniformly detected within this thin layer (and neglects other factors including, e.g., the solute orientation), the weighting function can be approximated by

$$\begin{aligned} w^{SFG}(z) &= 1 \text{ if } z \in [L/2 - l; L/2] \\ &= 0 \text{ otherwise} \end{aligned} \tag{S18}$$

X-ray photoelectron spectroscopy (XPS) probes the interfacial region over a depth determined by the photoelectron inelastic mean free path, and detects a weighted signal that exponentially decays with depth. The attenuation length λ depends on the incident photon energy and λ values in XPS experiments probing liquid interfaces have been suggested to be close to three molecular layers.[25,26] We thus consider a typical attenuation length of $\lambda = 9$ Å, and

$$\begin{aligned} w^{XPS}(z) &= e^{-d(z)/\lambda} \text{ if } z \in [-L/2; L/2] \\ &= 0 \text{ otherwise} \end{aligned} \tag{S19}$$

where $d(z)$ is the distance from position z to the nearest interface.

To model the OESI-MS experiments, we consider the acid fraction over a depth l that has been suggested to be on the order of a nanometer,

$$\begin{aligned} w^{OESI-MS}(z) &= 1 \text{ if } z \in [L/2 - l; L/2] \\ &= 0 \text{ otherwise} \end{aligned} \tag{S20}$$

System specific area. The impact of the system specific area, i.e. the surface/volume ratio of the aqueous system, can be examined by changing the thickness L of our slab system here, independently of the experimental probing depth. An increase in the specific area (i.e. a decrease in L) leads to a better stabilization of the released hydronium (a more negative $\left\langle \Delta G^H_{hyd} \right\rangle_L$ in eq. S8), and therefore an increase in the dissociation equilibrium constant $\langle K_A \rangle_L$ eq. S15 and a decrease in the average acid fraction x_A^L (eq. S17). To determine the value of L for a given slab, we consider the region of space where the water density is non-vanishing, which extends from the middle of the slab to approximately 1 Å above each GDS interface (see density profile in Fig. 2). While this definition is not unique, our choice allows to include all molecules that are adsorbed on the aqueous slab while not considering molecules that would be in the vapor phase.

Figure S9 shows how the measured acid/base ratio changes with the slab thickness L and with the experimental probing depth l (assuming a uniform sensitivity on this probing depth as in eqs. S18, S20 and relevant for, e.g., SFG and OESI-MS). It reveals that while the surface-specific SFG measurements are expected to report an increased acid/base ratio relative to the bulk for both nitric and formic acids, techniques with nanometric probing depth like OESI-MS employed on systems with a nanometric size can report acid/base ratios reduced relative to the bulk value. For formic acid, Fig. S9b shows that a reduced acid/base ratio and therefore an enhanced acidity would be measured for systems with $L/2 < 35$ Å and a probing depth $l > 10$ Å (For formic acid in a L =15.5 Å-thick slab with a l =7.75 Å probing depth, the calculated average dissociated fraction is approximately 5%; this fraction is smaller than the 14% dissociated fraction reported by OESI-MS experiments[27] for hexanoic acid in pH=2.1 microdroplets, but several important differences between our simulated formic acid system and the hexanoic acid experiments can affect the degree of dissociation, including the different pH conditions, different acid concentrations, and the different carboxylic acid alkyl chains). In contrast, for nitric acid, Fig. S9a shows that for the same system size and probing depth, an enhanced acid/ratio and thus a decreased acidity would be reported.

Figure S10 shows the acid/base ratios x_A^{XPS}/x_B^{XPS} expected with an XPS-like exponential weighting function eq. S19. When the slab thickness L increases, the acid/base ratio reaches a plateau value corresponding to an acid fraction $x_A^{XPS}/\left(x_A^{XPS}+x_B^{XPS}\right)$ for nitric acid of approximately 10 %, in excellent agreement with the 9 % value found in experiments conducted on micrometric systems.[26]

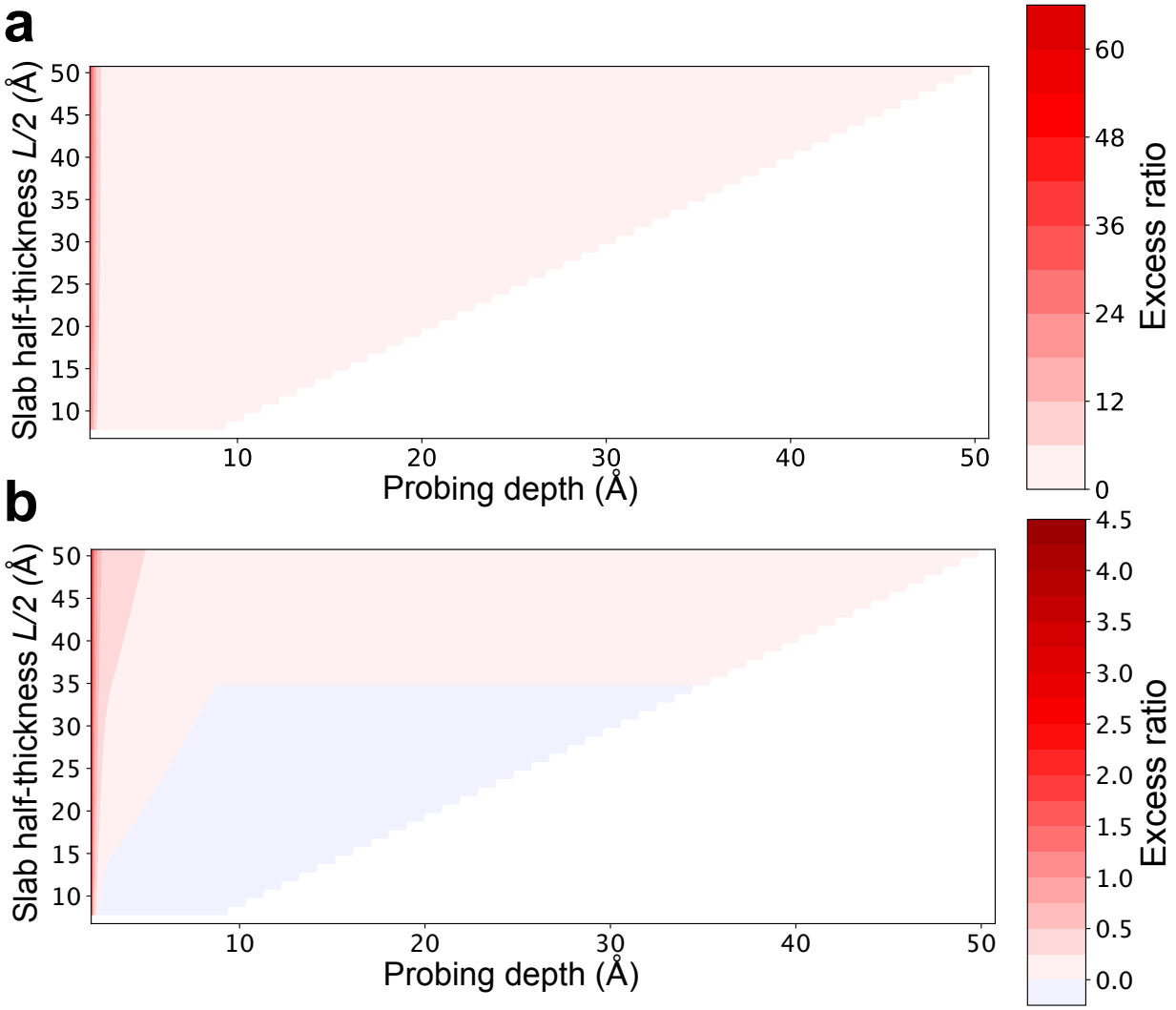


Figure S9: Excess acid/base ratio relative to the bulk $\frac{x_A/x_B}{x_A^{bulk}/x_B^{bulk}} - 1$ for a) nitric and b) formic acid for different probing depth l and slab thickness L (using weighting functions in eqs. S18, S20).

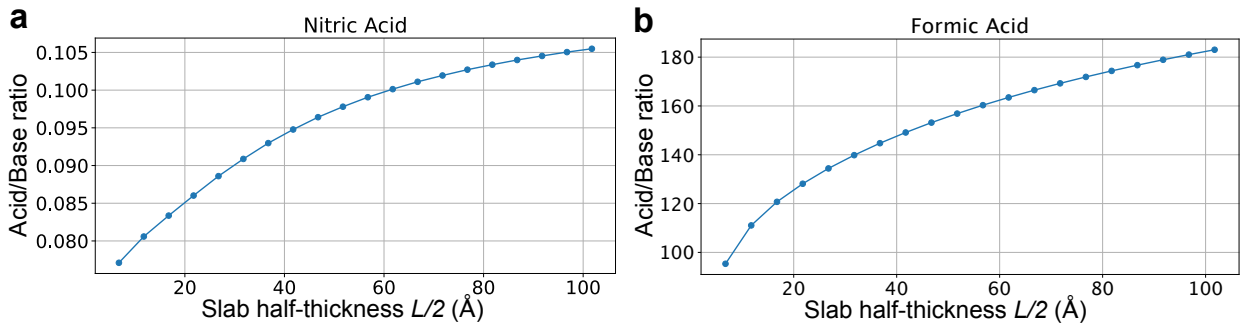


Figure S10: Acid/base ratios x_A^{XPS}/x_B^{XPS} of a) nitric and b) formic acids detected with an XPS-like exponential weighting function eq. S19 ($\lambda = 9$ Å) for increasing slab thickness L.

References

(1) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(2) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511–519.

(3) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol Phys.* **1984**, *52*, 255–268.

(4) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. **1985**, *31*, 1695–1697.

(5) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-Kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184.

(6) Zhang, L.; Han, J.; Wang, H.; Saidi, W. A.; Car, R.; E, W. End-To-End Symmetry Preserving Inter-Atomic Potential Energy Model for Finite and Extended Systems. *Adv. Neural. Inf. Process. Syst.* **2018**, 4436–4446.

(7) Yue, S.; Muniz, M. C.; Calegari Andrade, M. F.; Zhang, L.; Car, R.; Panagiotopoulos, A. Z. When Do Short-Range Atomistic Machine-Learning Models Fall Short? *J. Chem. Phys.* **2021**, *154*, 034111.

(8) Elstner, M.; Seifert, G. Density Functional Tight Binding. *Philos. Trans. A. Math. Phys. Eng. Sci.* **2014**, *372*, 20120483.

(9) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B* **1998**, *58*, 7260–7268.