Deprotonation of formic acid in collisions with a liquid water surface studied by molecular dynamics and metadynamics simulations: Electronic supplementary information (ESI)

Garold Murdachaew,*,[†] Gilbert M. Nathanson,[‡] R. Benny Gerber,^{†,¶,§} and Lauri

 $Halonen^{\dagger}$

Laboratory of Physical Chemistry, Department of Chemistry, P.O. Box 55 (A.I. Virtasen aukio 1), FI-00014 University of Helsinki, Finland, Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States, Institute of Chemistry and the Fritz Haber Research Center, The Hebrew University, 91904 Jerusalem, Israel, and Department of Chemistry, University of California, Irvine, California 92697, United States

E-mail: garold.murdachaew@helsinki.fi

October 17, 2016

^{*}To whom correspondence should be addressed

[†]Laboratory of Physical Chemistry, Department of Chemistry, P.O. Box 55 (A.I. Virtasen aukio 1), FI-00014 University of Helsinki, Finland

[‡]Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States ¶Institute of Chemistry and the Fritz Haber Research Center, The Hebrew University, 91904 Jerusalem, Israel

[§]Department of Chemistry, University of California, Irvine, California 92697, United States

1 Additional validation of theory level

As discussed in the main text, BLYP-D2 has been shown to correctly reproduce the density of bulk liquid water. We confirmed that the level of theory used for the majority of the calculations, BLYP-D2/DZVP, correctly reproduced the density and H-bond populations of the water molecules in the slab. We also ensured that the radial distribution functions (RDFs) in the water slab (interior region) and in the bulk water agree reasonably with one another and that the latter agree well with experiment. This was described in our recent work on HCl scattering from water using the identical setup, see ref. 1 and its ESI.

Thus, here we restrict ourselves to confirming that BLYP-D2/DZVP correctly describes the geometries of H₂O, *trans*-HCOOH, and the *trans*-HCOOH····H₂O minimum energy complex (designated structure FAZ1 in ref. 2). Fig. S1 shows that with all basis sets, the complex assumes a nonplanar geometry, with the free OH of the water pointing out of the plane established by the formic acid (FA) or the heavy atoms of the complex. Two hydrogen bonds are present. This is also in good agreement with the B3LYP² and MP2 benchmarks,³ as well as with the structure obtained experimentally.²

In Table S1 and Fig. S1 it is seen that the monomer and dimer geometries are reasonably reproduced by BLYP-D2, even with the DZVP basis set, although DZVP-MOLOPT-SR and TZV2P give slightly better converged results for the geometries. Comparing BLYP-D2/DZVP to BLYP-D2/TZV2P, the average differences in intramolecular bond lengths and bond angles are ~ 0.01 Å and ~ 1°, respectively. The differences in the intermolecular parameters are sometimes larger, being in one instance ~ 0.02 Å and ~ 2°. Comparing to the B3LYP² and MP2 benchmarks,³ the agreement is also reasonable.

With regard to the interaction or binding energy $(-D_e)$ of the complex, it is seen by comparing BLYP/TZV2P and BLYP-D2/TZV2P that the dispersion correction deepens it by about 20%. The dispersion correction has negligible effects on the structures of the monomers or the dimer. However, as mentioned, it is critical for correct description of liquid water, and thus it is necessary to retain it. For clusters, the only significant issue with use of the DZVP basis set is that it does lead to a deepening of about 17% in the interaction energy (compare interaction energy obtained with BLYP-D2/DZVP to that obtained with BLYP-D2/DZVP-MOLOPT-SR or BLYP-D2/TZV2P; the deepening is mainly due to basis set superposition error—BSSE). However, in this work we are mainly interested in forces for molecular dynamics and the dimer structures in Fig. S1 and Table S1 are nearly identical, indicating that the forces are well-represented with the smaller basis set. Thus we consider DZVP a reasonable compromise given the computational costs of calculating many molecular dynamics trajectories for large systems.

Table S1: Theory level $HCOOH \cdots H_2O$ complex (to two literature benchman	valıdatıon: (FAZ1) and th rks	Geometric p e interaction	arameters of H ₂ O energy of the latter	, <i>trans</i> -HC(as calculate	JOH, and d with diffe	the minimum rent levels of the	energy <i>trans</i> - eory compared
molecule		BLYP-D2/ DZVP	BLYP-D2/ DZVP-MOLOPT-SR	BLYP-D2/ TZV2P	BLYP/ TZV2P	$\begin{array}{c} \mathrm{B3LYP}/\\ \mathrm{6-311++G(d,p)}\end{array}$	$\frac{MP2(full)}{6-311++G(3df,2p)}$
reference	Comotnio	this work ^a	this work ^{a} $\int_{a}^{a} \int_{a}^{a} \int_{a}^{a}$	this work ^a	this work ^a	ref. 2	ref. 3
	CEOILIEUTIC	: parameters (no	und tenguns are m A ar	ia pona angles	are III uegre		
Isolated H_2O	Ow-Hw	0.98	0.97	0.97	0.97	0.962	0.9577
	Hw-Ow-Hw	102.8	104.4	104.3	104.2	105.0	104.59
Isolated <i>trans</i> -HCOOH	H-O	0.99	0.98	0.98	0.98	0.971	0.9673
	C=0	1.22	1.22	1.21	1.21	1.199	1.2032
	C-H	1.12	1.10	1.11	1.11	1.098	1.0897
	C-0	1.37	1.38	1.37	1.37	1.346	1.3452
	H-O-C	106.6	107.2	107.1	106.9	108.0	106.92
	0-C=0	125.3	125.1	125.3	125.3	125.2	125.17
	H-C=O	125.7	126.0	125.6	125.6	125.2	125.18
trans-HCOOH····H ₂ O: FAZ1	Ow-H1w	1.00	0.99	0.99	0.99		0.969
1	Ow-H2w	0.98	0.97	0.97	0.97		0.956
	Hw-Ow-Hw	104.4	106.0	105.6	105.8		
	H-O	1.02	1.01	1.01	1.01	0.989	0.985
	C=0	1.24	1.23	1.23	1.23	1.211	1.211
	C-H	1.12	1.10	1.11	1.10	1.098	1.092
	C-0	1.35	1.35	1.35	1.35	1.328	1.324
	Н-О-С	107.4	108.1	107.8	107.6		107.8
	0-C=0	126.1	126.0	126.1	125.9		125.9
	H-C=O	123.2	123.6	123.3	123.4		
	$O)H\cdots O_{M}$	1.73	1.79	1.76	1.77		1.779
	$(C=)O\cdots H1w$	2.01	1.95	2.03	2.03		2.025
	$\mathrm{O}\text{-}\mathrm{H}\text{-}\text{\cdot}\mathrm{O}_{\mathrm{W}}$	158.2	155.2	157.8	158.2		157.7
	$H \cdots O_{W} - H1_{W}$	86.7	84.7	86.0	85.5		86.6
	$C-0\cdots H1_W$	104.9	105.7	105.2	105.5		106.4
	$O\cdots H1$ w- O w	136.8	140.3	137.1	137.2		135.2
Interac	ction energy (in l	scal/mol; results	s with counterpoise cor	rection—CP—	of BSSE are	in parentheses)	
<i>trans</i> -HCOOH····H ₂ O: FAZ1	$-D_e$	-13.9	-11.9	-11.8	-9.6	-10.3	-10.7
							(-9.63)

^a All geometry optimizations were performed with CP2K and the indicated DFT functional and basis set. The system was a non-periodic box with a side length of 30 Å.

(a) BLYP-D2/DZVP

(b) BLYP-D2/DZVP-MOLOPT-SR



Figure S1: Geometries of the minimum energy *trans*-HCOOH····H₂O complex (FAZ1) as obtained using BLYP or BLYP-D2 with various basis sets. See Table S1 and text for details.

2 Scattering simulations of formic acid at the air-water interface: Additional results

2.1 All scattering trajectories: Formic acid height above water surface and hydroxyl bond length

As discussed in the main text, we calculated at the simulation temperature of 300 K a total of 45 scattering trajectories with durations of 8–50 ps. Of these, 37 were thermal scattering/gentle landing trajectories, and 8 were trajectories with added kinetic energies nk_BT , n = 1, 2, 4, 10 (2.5, 5.0, 10.0, or 24.9 kJ/mol, respectively), directed downward at 45° to the surface normal.

The main text discussed and presented 6 trajectories, all of the thermal scattering/gentle landing type. Figs. S2 and S3 present the FA height above the water surface z and its hydroxyl O-H bond length $r_{\rm OH}$ for all 45 trajectories. In each figure, panel (a) shows the initial 10 ps to allow a better view of initial events while panel (b) presents the full simulation time. As observed in the main text, adsorption is seen exclusively over this relatively short time period, and trajectories 1 and 2 are seen to dissociate. However, many other trajectories do experience stretched O-H bonds.

In general, the trajectories with added kinetic energy, marked nkT(a,b), n = 1, 2, 4, 10, do not behave differently from the trajectories with only thermal collision energy with regard to deprotonation, as seen in Fig. S3. This is unlike the case of the strong acid HCl colliding with the water surface, where increasing its collision energy generally resulted in faster deprotonation.¹ In Fig. S2 the trajectories with added kinetic energy do require 2–3 ps longer to stabilize their height z at the air-water interface. Further studies of the equilibration of such collisions is the subject of future work, and is beyond the scope of this paper.



The vertical position of FA's carbon atom relative to the center of the water slab is shown for all 45 collision trajectories. The black lines mark the position of the water slab's Gibbs dividing surface z_{GDS} and the surface's interfacial width $\pm 2\delta$. The colored dotted lines (blue, center of the slab; red, subsurface region; and green, surface region, see Fig. 2 in the main text) mark the vertical positions at which FA was fixed in the metadynamics simulations and are drawn here to show how often the unrestrained trajectories visit these regions. Figure S2:





3 Equilibrium simulations of formic acid in aqueous environments: Additional results

3.1 Formic acid hydroxyl bond length

Fig. S4 shows the FA hydroxyl bond length r_{OH} from the equilibrium simulations (production portions) of FA in the four aqueous environments. FA in the subsurface is observed to spontaneously deprotonate.

3.2 Formic acid H-bonds with water

Fig. S5 shows time series of the number of H-bonds that FA forms with water molecules for the four aqueous environments. These data were obtained from equilibrium simulations. Table 1 in the main text summarizes the H-bond results in terms of averages.

Spontaneous deprotonation was observed when FA was placed in the subsurface. The arrows labeled b–f refer to the snapshots shown in Fig. S6.

3.3 Formic acid spontaneous deprotonation observed in the subsurface equilibrium calculation

The panels in Fig. S6 show snapshots from the equilibrium simulation of FA constrained at the subsurface. Spontaneous deprotonation and PT is seen.

3.4 Formic acid orientation at the air-water interface

The distributions for the remaining two orientation angles we considered (see Fig. 1 of the main text for definitions of the orientation angles) are presented in Fig. S7. These are not independent from the two presented in the main text. As the FA transitions from the surface into the subsurface, it rotates from an initial nearly vertical orientation to a leaning

orientation. This is most clearly seen in the change of the angle between the plane of the FA molecule and the z-axis, θ_n , which changes from about 97° to about 27°.



Figure S4: The O-H separation in FA is shown is shown for equilibrium simulations in the four aqueous environments (bulk water, and center, subsurface, and surface of the water slab). The dashed black lines mark the approximate O-H separations for various states of the aqueous FA.



at 4 H-bonds in the rightmost columns indicate the approximate H-bonding threshold for FA ionization (3-4 H-bonds). The Figure S5: The number of H-bonds that FA forms with water molecules is shown for the four aqueous environments. These data were obtained from equilibrium simulations. The rows, top to bottom, show FA in bulk liquid; and constrained in the water slab the FA hydroxyl group (-OH), the FA carbonyl group (-C=O), and the total number of H-bonds, respectively. The dotted line at the center, at the subsurface, and at the surface. Left to right, each row shows the number of FA-water H-bonds involving arrows labeled b-f refer to the snapshots shown in Fig. S6. See also Table 1 in the main text for a summary of the average numbers of FA-water H-bonds in each environment.



at start











(e) t = 13.95 ps: SSIP2 (with Zundel cation)









Figure S6: Snapshots (top view) from the equilibrium simulation of FA constrained at the subsurface showing deprotonation and PT. See also Fig. 5 of the main text for a description of image details.



(a) $\cos \theta_n$ and (b) $\cos \theta_{CO}$. The extracted average values and standard deviations of the angles here defined using $\theta \equiv \cos^{-1} \langle \cos \theta \rangle$ are: (a) $\theta_{n}^{surface} = 97 \pm 24^{\circ}$, $\theta_{n}^{subsurface} = 27 \pm 14^{\circ}$; and (b) $\theta_{CO}^{surface} = 81 \pm 10^{\circ}$, Figure S7: Distributions of the cosines of two orientation angles of neutral trans-FA at two depths in the interfacial region $\theta_{\rm CO}^{\rm subsurface} = 84 \pm 12^{\circ}$. See the main text for the other two orientation angles. obtained from equilibrium simulations:

4 Metadynamics simulations: Background and additional details

In order to better explore the deprotonation mechanisms of FA in these aqueous systems and obtain the free energy of dissociation barriers, an enhanced sampling method was required. We employed metadynamics,^{4,5} in particular, well-tempered metadynamics (WTmetaD),⁶ which has been shown to converge asymptotically.⁷ See also ref. 8 for a recent review. The WTMetaD simulations always began with an equilibrated structure (with neutral, *trans*-FA). As in ref. 9, we examined the activity of the interface by constraining the C atom of FA at two heights in the interfacial region, at the subsurface ($z = z_{\text{GDS}} - 1$ Å = 3.45 Å), at the subsurface ($z = z_{\text{GDS}} + 1$ Å = 5.45 Å), and at the center of the slab (z = 0 Å). The last result can be compared to the analogous calculation where FA is solvated in the bulk liquid water with 63 water molecules. After an equilibration in the *NVT* ensemble, the WTMetaD calculations were performed in the same ensemble. Thus, the free energy differences are properly Helmholtz free energy differences ΔF , but, with our slab setup, where the internal pressure will be approximately constant, they approximate Gibbs free energy differences ΔG . For the case of the liquid water simulation, this is even more approximate.

Additional information on the selection of collective variables and on both metadynamics and well-tempered metadynamics can be found in refs. 4,6,8,10–12 and in the PLUMED 2 package and documentation.^{13–15} Briefly, in metadynamics, the configurational space is explored with a total potential

$$U(\mathbf{r}, \mathbf{s}, t) = U_0(\mathbf{r}) + V(\mathbf{s}, t), \tag{1}$$

where U_0 is the unbiased interatomic potential, V is the applied bias potential, $\mathbf{r} = (x_1, y_1, z_1, \dots, x_N, y_N, z_N)$ is the full set of nuclear coordinates within a system of N atoms, and the chosen set of metadynamics variables that one wishes to sample is in general a vector of

dimension d of collective variables (CVs) dependent on the nuclear coordinates (it may be explicitly dependent on only a subset thereof), $\mathbf{s} = \mathbf{s}(\mathbf{r})$, $\mathbf{s} = (s_1, s_2, \ldots, s_d)$. In practice, dis no more than 3 and usually 2 or even 1. The bias potential is time-dependent (i.e., its history is recorded) to ensure that previously visited regions of CV phase space (\mathbf{s}') at prior times (t') are avoided, thus ensuring more efficient sampling. This is done by depositing small repulsive Gaussians centered at \mathbf{s}' . Thus, the form of the cumulative bias potential is

$$V(\mathbf{s},t) = \sum_{t'=\tau_G, 2\tau_G, \dots}^{t' < t} W \exp\left(-\sum_{i=1}^d \frac{(s_i - s'_i)^2}{2\sigma_i^2}\right),$$
(2)

where τ_G is the time interval between depositing Gaussians, $\mathbf{s} = \mathbf{s}(\mathbf{r}(t))$ is the instantaneous value of the CVs, $\mathbf{s}' = \mathbf{s}'(\mathbf{r}(t'))$ are previously visited positions which are the centers for deposited Gaussians, $\boldsymbol{\sigma}$ is the vector of Gaussian widths corresponding to \mathbf{s} , and W is the constant height of the Gaussian.

In well-tempered metadynamics (WTMetaD), overfilling of basins in the multidimensional energy surface is prevented and sampling of the CV space enhanced by making the height W history dependent as well,

$$W = W(t') = W_0 \exp\left(-\frac{V(\mathbf{s}', t')}{k_B \Delta T}\right),\tag{3}$$

where ΔT is chosen to be a few times the system temperature T. This ensures that the sampling of **s** is performed at the higher temperature $T + \Delta T$ (one speaks of the acceleration factor, γ , $\gamma = \frac{T + \Delta T}{T}$). Note that the dependence of W on $V(\mathbf{s}, t - \tau_G)$ implies that the Gaussian height depends on the cumulative bias of the previous step, thus avoiding a circular definition. As the run proceeds, the incremental bias approaches zero and the total bias approaches a constant value.⁷

The usefulness of metadynamics is that in the long time limit, it converges to the free-

energy surface (plus an additive constant),

$$V(\mathbf{s}, t \to \infty) = -F(\mathbf{s}) + C. \tag{4}$$

Similarly, for WTMetaD,

$$V(\mathbf{s}, t \to \infty) = -\frac{\Delta T}{T + \Delta T} F(\mathbf{s}) + C.$$
(5)

In WTMetaD, the parameter ΔT controls the free-energy surface exploration. The limits are clear: $\Delta T = 0$ corresponds to standard molecular dynamics, and $\Delta T \rightarrow \infty$ to standard metadynamics.

After some experimentation and tuning, the following parameters (which are typically found also in refs. 6,7) were selected. A single CV, $\mathbf{s} \equiv s = n_{\text{OH}}$, which measures the coordination of the hydroxyl oxygen of FA by its H atom was found to be relevant,

$$n_{\rm OH} = \frac{1 - \left(\frac{r_{\rm OH}}{r_0}\right)^n}{1 - \left(\frac{r_{\rm OH}}{r_0}\right)^m},\tag{6}$$

where r_{OH} is the instantaneous FA hydroxyl bond length. Use of coordination number rather than bond length as a CV has been shown to be more robust for exploring aqueous deprotonation and proton transfer (PT), particularly in stabilizing and better sampling shortlived states such as the CIP and SSIP.¹⁶ The formula eqn 6 and the parameters $r_0 = 1.6$ Å, n = 6, m = 12 are typical, and are chosen so that as the system transitions from neutral FA to ionized FA (CIP and SSIP), n_{OH} varies smoothly from about 1 (in practice about 0.9) to about 0 (in practice about 0.3).

The other parameters were also typical and were likewise tested before use. The deposition rate for Gaussian hills was 250 steps ($\tau_G = 125$ fs), the Gaussian width selected for sampling the space in a reasonable manner was $\boldsymbol{\sigma} \equiv \boldsymbol{\sigma} = 0.03$ (approximately equal to the size of the fluctuations in an unbiased simulation, as observed by analyzing the preceding unbiased NVT trajectory using PLUMED), and the initial Gaussian height was $W_0 = 5 \times 10^{-4}$ Hartree = 1.3128 kJ/mol = 0.3138 kcal/mol (approximately 0.1 of the FA deprotonation barrier height). We chose $\Delta T = 1200$ K (CV temperature of 1500 K and $\gamma = 5$). The WT-MetaD simulations were run for 14–18 ps. We confirmed good sampling and convergence of the relevant portions of $n_{\rm OH}$ and of the relevant free energy differences.

5 Metadynamics simulations of formic acid in aqueous environments: Additional results

5.1 Formic acid H-bonds with water

Fig. S8 shows the number of H-bonds that FA forms with nearby water molecules in the metadynamics trajectories in the four aqueous environments. It is seen that the number of H-bonds increase as formate is formed. The arrows labeled b–d for the bulk water and subsurface metadynamics simulations refer to the snapshots shown in Fig. 11 of the main text.

5.2 Formic acid free energy of deprotonation

Fig. S9(a-b) shows the sampling of the four metadynamics simulations in the CV, n_{OH} , and also in the r_{OH} coordinate. The initial barrier (states a and TS_{ab}, corresponding to the initial neutral *trans*-FA and the initial barrier to the shared proton state, respectively, see the main text for details) are well-converged while the other states (b, c, and d, corresponding to the shared proton, CIP, and SSIP states) are more approximate in some of the four simulations. See Fig. 11 in the main text for the obtained deprotonation free energy profiles.



Figure S8: The number of H-bonds that FA forms with water molecules. These data were obtained from the metadynamics H-bonds in the rightmost columns indicate the approximate H-bonding threshold for FA ionization (3–4 H-bonds). The arrows simulations in various environments. The rows, top to bottom, show FA in bulk liquid; constrained at the center of the water slab; at the subsurface, and at the surface. Left to right, each row shows the number of FA-water H-bonds involving the FA hydroxyl group (-OH), the FA carbonyl group (-C=O), and the total number of H-bonds, respectively. The dotted line at 4 abeled b-d for the bulk water and subsurface metadynamics simulations refer to the snapshots shown in Fig. 11 of the main text. (b) Sampling: shown in the equivalent $r_{\rm OH}$ bond length

(a) Sampling: shown in $n_{\rm OH}~{\rm CV}$



Figure S9: Monitoring the sampling in the well-tempered metadynamics simulations of FA deprotonation in liquid water and at three depths in the water slab. See text for details.

References

- Partanen, L.; Murdachaew, G.; Gerber, R. B.; Halonen, L. Temperature and collision energy effects on dissociation of hydrochloric acid on water surfaces. *Phys. Chem. Chem. Phys.* 2016, 18, 13432–13442.
- (2) Aloisio, S.; Hintze, P. E.; Vaida, V. The Hydration of Formic Acid. J. Phys. Chem. A 2002, 106, 363–370.
- (3) Priem, D.; Ha, T.-K.; Bauder, A. Rotational spectra and structures of three hydrogen-bonded complexes between formic acid and water. J. Chem. Phys. 2000, 113, 169–175.
- (4) Laio, A.; Parrinello, M. Escaping free-energy minima. Proc. Natl. Acad. Sci. USA (PNAS) 2002, 99, 12562–12566.
- (5) Iannuzzi, M.; Laio, A.; Parrinello, M. Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics. *Phys. Rev. Lett.* **2003**, *90*, 238302.
- (6) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. Phys. Rev. Lett. 2008, 100, 020603.
- (7) Dama, J. F.; Parrinello, M.; Voth, G. A. Well-Tempered Metadynamics Converges Asymptotically. *Phys. Rev. Lett.* **2014**, *112*, 240602.
- (8) Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. Annu. Rev. Phys. Chem. 2016, 67, 159–184.
- (9) 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.
- (10) Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.* 2008, 71, 126601.
- (11) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. WIREs Comput. Mol. Sci. 2011, 1, 826–843.
- (12) Sutto, L.; Marsili, S.; Gervasio, F. L. New advances in metadynamics. WIREs Comput. Mol. Sci. 2012, 2, 771–779.
- (13) PLUMED 2, http://www.plumed.org/, version 2.2.2, 2016.

- (14) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961–1972.
- (15) PLUMED 2: Documentation, http://plumed.github.io/doc-v2.2/user-doc/html/index.html, version 2.2.2, 2016.
- (16) Sprik, M. Computation of the pK of liquid water using coordination constraints. Chem. Phys. 2000, 258, 139–150.