Supporting Information for

Molecular sinkers: X-ray photoemission and atomistic simulations of benzoic acid and benzoate at the aqueous solution/vapor interface

Niklas Ottosson[1,a], Anastasia O. Romanova[2,b], Johan Söderström[a], Olle Björneholm[a], Gunnar Öhrwall[c], and Maxim V. Fedorov[2,b]

[a]*Department of Physics and Astronomy, Uppsala University, SE-751 20 Uppsala, Sweden*
[b]*Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, D 04103, Leipzig, Germany*
[c]*MAX-lab, Lund University, Box 118, SE-221 00 Lund, Sweden*

Content

---

[1] Present address: FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam

[2] Present address: Department of Physics, Strathclyde University, John Anderson Building, 107 Rottenrow East, Glasgow, U.K. G4 0NG

Simulation details

The simulation cells were prepared as follows: Molecular structures (20 molecular benzoic acids or 87 benzoates) were randomly placed in a periodic cell of size 4.1x4.1x10.0 nm. Solutes were solvated with water (TIP4P). Water molecules overlapping with any solute atom were removed. In the case of benzoate molecules, the system was neutralized by random replacing of water molecules with sodium ions (87 Na^+).

After the NPT equilibration run (1 ns) of the simulations, a vacuum slab of 7 (6+1) nm was added to the equilibrated system; the resulting geometry of the simulation cell for the productive run is shown on Figure S1 that represents a snapshot from the productive run. We performed productive simulations (55ns) in the NVT ensemble at constant temperature (T=300K) using the Berendsen thermostat[1] with a relaxation time equal to 1ps.

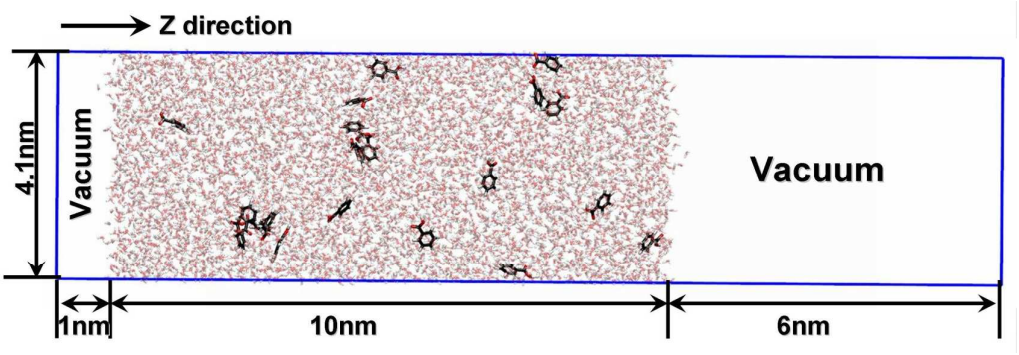

Figure S1. Geometry of the simulation cell shown in the XZ plane.

Table S1. Force-field atom parameters used in the molecular dynamic simulations. σ and ε are the distance and energy parameters of the Lennard-Jones interaction potential; σ – is the distance at which the inter-particle potential is zero; ε is the depth of the potential well. q is the partial atom charge. Geometrical OPLS mixing rules were used for the Lennard-Jones interaction parameters between different atoms.

Atom	**σ, Å**	**ε, kJ mol^{-1}**	**q, $\bar{e}$**
Molecular benzoic acid			
C4a	3.55	0.29288	-0.12
H6a	2.42	0.12552	0.12
C2a	3.75	0.43932	0.64
O4a	3.00	0.71128	-0.53
O2a	2.96	0.87864	-0.44
H8a	0.50	0.12552	0.45
Ionized benzoic acid (benzoate)			
C4a	3.55	0.29288	-0.12
H6a	2.42	0.12552	0.12
C2a	3.75	0.43932	0.71
O2a	2.96	0.87864	-0.80
Water			
OW	3.16	0.64852	0.00
HW	0.00	0.000	0.52
PSEUDO(MW)	0.00	0.000	-1.04
Ions			
Na^+	4.070	0.20930	1

During the productive run we controlled the change of the molecular density at the interfaces and in the bulk. The change of the molecular density during the productive run for molecular benzoic acid and benzoate are represented at Figures S2a and S2b respectively.

a) b)

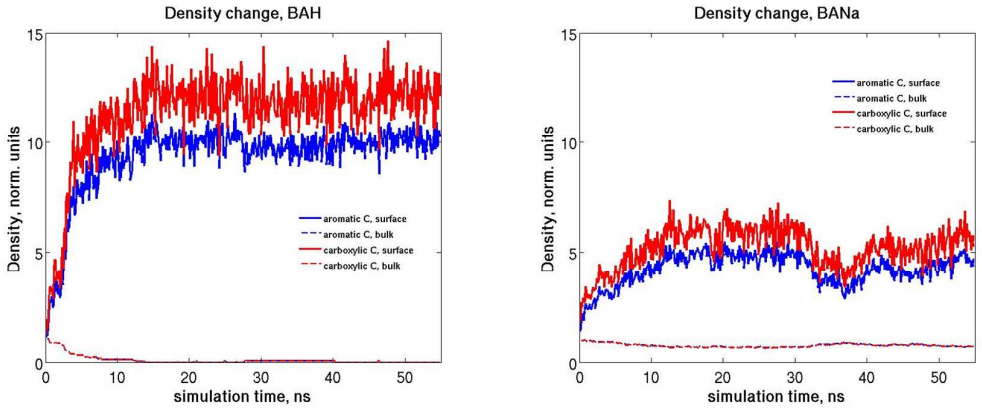

Figure S2. Changes of molecular densities at the interface and in the bulk during the simulations; a) benzoic acid; b) benzoate. We note that the average density of BAH in the bulk between 15 ns and 55 ns does not go to exact zero; the average value of the BAH density in the bulk is 0.065 normalized units that corresponds to the concentration 0.013 M; that is in a qualitative agreement with the maximum solubility of BAH at room temperature (according to [8], experimental value of the solubility limit for BAH in water at 25 C is 0.02 M).

In our simulations the concentration of the neutral benzoic acid was 0.2M (according to [8], solubility of the neutral benzoic acid in water at 25 C is 0.02M). There were two main reasons to do so. Firstly, collection of a sufficient statistics with the 0.02M solution of the benzoic acid would require an enormous simulation time. Second, according to the experimental results (see main text, Fig. 1) benzoic acid is strongly surface segregated, meaning that the concentration of benzoic acid in the surface layer is much higher than the analytical solubility. In our simulations the surface area available for one molecule is at least 4 times the surface area required for each benzoic acid molecule. Hence we believe that the surface behavior of the molecules should be similar to that at realistic concentrations. To check this hypothesis we performed a short simulation of 0.02M benzoic acid. Although very noisy, the system in this simulation shows the same tendency as the system reported in the present paper. In addition, we performed a free energy analysis of the interfacial transfer of one BAH molecule from the vapor phase to the bulk by calculation of the Potential of Mean Force for the BAH molecule travelling along the z-direction through the interface (these calculations are discussed in details below in a separate section at the end of the document). These calculations also shown that although the hydration free en-

ergy of the BAH in the bulk is negative it preferentially adsorbs at the vapor/water interface.

Angle distributions

A rough estimation of the molecular surface orientation can be performed if the molecular density profiles are known. The comparison of the molecular linear sizes with the positions of the maxima of the density profiles allows us to estimate the molecule surface orientation. However, this approach has a number of limitations. First, we could only obtain information about one angle (between the main molecular axes and the water surface). Second, we would have to be sure that the particle densities are simple peaks and do not consist of a number of sub-peaks corresponding to distinctly different, co-existing molecular orientations. We have therefore undertaken a direct analysis of the angles at each simulation step.

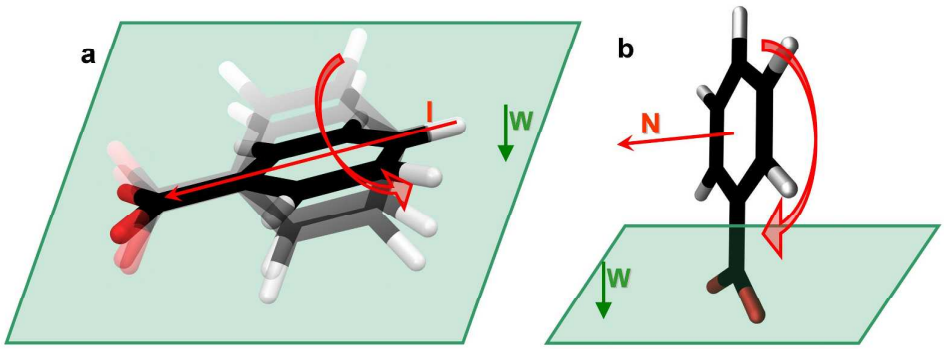

Figure S3. Possible unregistered movements of the benzoic acid molecule. a) The angle *main molecular axis – surface* is known and fixed; b) The *angle ring plane – surface* is known and fixed.

It should be clear that we cannot describe the orientation of the benzoic acid molecule fully with a single angle. To describe the orientation of BAH/BA^- molecule we used a set of two angles. The angle definitions are represented in Fig. 2d in the main text. If we use only $\Theta_{I\text{-}W}$ we cannot distinguish a molecule with the aromatic ring situated flat on the water surface and a molecule with the aromatic ring plane rotated to the surface if the main molecular axis is parallel to the surface. However, we can clearly see the orientation of the COOH/COO^- group of the molecule (see Fig.

S3a). If we use only $\Phi_{N\text{-}W}$ we can not distinguish a molecule with the $COOH/COO^-$ pointing to water and a molecule with the $COOH/COO^-$ pointing to any other direction. However we can clearly see the orientation of the plane of the aromatic ring (see Fig. S3b). Combining these two angles completely describes the molecular orientation.

Jacobian determinant of the angular transformation

Vectors used for the angular distribution analysis are orthogonal to each other, meaning that they cannot take any random orientation with respect to each other. Let vector **I** form an angle **Θ** with vector **W**. In this case vector **N** forms an angle **Φ** with the vector **W** such that $(\mathbf{\Theta} - 90^o) < \mathbf{\Phi} < (\mathbf{\Theta} + 90^o)$. In other words, for every given value of one angle (**Θ** or **Φ**) we have a given set of values for the other angle (**Φ** or **Θ**). Applying the formula for conditioned probabilities we have an additional member to consider in the Jacobian determinant of the angular transformation: $\frac{d\Phi}{dt} = \frac{\sqrt{1-\cos^2\Phi-\cos^2\Theta}}{\sin\Phi}$, where *t* is a parameter for the conditioned set of angles. Implementing this formula to Jacobian determinant of the angular transformation we have finally: $J = 4\pi\frac{\sqrt{1-\cos^2\Theta-\cos^2\Phi}}{\sin\Theta\sin\Phi}$.

Calculation of ring/carboxylic C1s PE ratios from the MD density profiles

To make a more direct connection between the simulations and the PE experiments we have employed a simple attenuation model, as developed in Ref. [2], to simulate the expected PE intensity ratios from the MD density profiles. Below we shortly discuss the assumptions made in this model and present the results thereof when applied to the current problem.

PE signals can generally be expressed as the integral over the density function, $\rho_A(z)$, associated with a molecular or atomic species *A*, exponentially attenuated with the electron inelastic mean free path λ. The PE emission density *I(z)* from a depth *z* in the sample is proportional to the density of the emitting species. Since λ, to a good

approximation, is inversely proportional to the integrated particle density (from an emission depth z to 0) [2], the total recorded photoelectron signal I from a species A, can be expressed as

$$I = \alpha F \sigma T \int_0^\infty \rho_A(z) \cdot \exp\left(-\frac{1}{\lambda}\int_0^z \frac{\rho_{tot}(y)}{\rho_0}\, dy\right) dz \qquad \text{(eq. 1)}$$

where ρ_{tot} is the sum of densities of all species and ρ_0 is the average density in the bulk, i.e., $\rho_{tot}(\infty) = \rho_0$. Further, σ is the photon energy dependent photoionization cross-section, F is the total photon flux, α is an experimental alignment factor, and T is the transmission function of the electron analyzer.

Given the obtained density functions of the respective ring and carboxylic/carboxylate carbons from the MD simulations (shown in Fig. 2a and b in the main text), the calculation of the associated photoemission signal intensities is straightforward after setting the pre-factors in eq. 1 to unity – we will below return to what is thereby implicitly assumed. Panels a) and b) of SI Fig. S4 show the simulated photoemission signal intensities from the aromatic and carboxylic/carboxylate carbon in the case of the simulated BAH and BA^-, respectively, given as function of the inelastic mean free path length λ. Furthermore, panels c) and d) show the ratio of the respective ring and carboxylic signal as function of the same parameter. This enables an explicit comparison between the experimental PE ratios and the simulated MD profiles. For large values of λ we see that we indeed approach the stoichiometric ring/carboxyl ratio of 6. For shorter attenuation lengths the values increase, more steeply for BA^- and less steeply for BAH, in excellent agreement with the experimental results.

To be able to make a fully quantitative comparison between the experimental and simulated PE ratios, the energy dependent inelastic attenuation function for photoemission from the aqueous phase must first be known. Unfortunately, this is not the case at present and while an attempt to derive this quantity was made in Ref. [2] the value for a specific aqueous solution at a given kinetic energy must be considered unknown. At present, we believe that the effective attenuation length from a liquid micro-jet near 70 eV must be between a half and one nanometer. We thus conclude that too small values of λ (<0.2nm) have to be assumed to accurately reproduce the experimental ring/carboxyl PE ratios (7.8/10.1 for BAH/BA^-). This could suggest that the

simulations are slightly overemphasizing the nearly flat-lying conformers. However, the relative change in the PE ratio upon going from BAH to BA^- appear to be in almost quantitative agreement.

Let us conclude by considering the pre-factors in equation 1 and how they might influence the experimental results. Since the carboxylic and ring carbon signals are obtained from the same spectrum, the pre-factors α and F cancels when forming the intensity ratio of the two peaks. Given that the kinetic energies in the current measurements lies around 70 eV, and the chemical shift between the two species compared is only ~4 eV, the transmission term also cancels to a very good approximation. The photoionization cross section σ for the two types of carbons are also similar, but by directly comparing the PE signal deviation from stoichiometric proportions we implicitly assume that they are identical, which is not necessarily always the case. Photon energy dependent variations in the cross section for C1s core-level photoemission of small halogenated organic molecules have however been found to extend surprisingly far above the ionization threshold [4], which might re-scale the experimentally ratios to some extent, relative to what would have been obtained under full cancellation of σ. However, as discussed in the main text, the important experimental observable here is the *change* in the ring/COO(H) C1s PE ratio when going from the neutral acid to benzoate. Therefore, the comparison should be unaffected by the phenomenon discussed in reference [4].

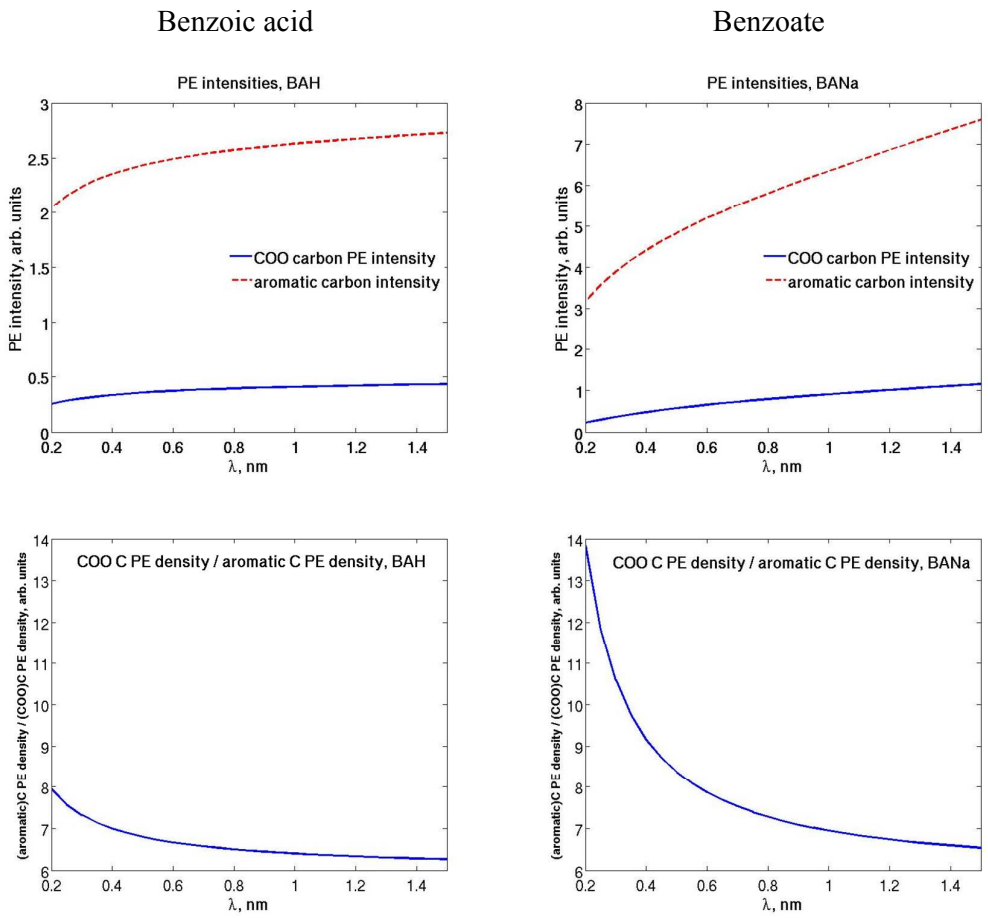

Figure S4. Panels a) and b) show the calculated signal intensity from the ring and carboxylic/carboxylate carbons from the MD simulated density profiles of the BAH and BA^- solutions, respectively, as function of the inelastic mean free path length λ. Panels c) and d) show the corresponding ring/carboxyl PE ratios.

Calculations of the Potential of Mean Force for one BAH molecule along the z-direction: free energy profile of BAH travelling through the water slab

The density profiles of neutral benzoic acid show a very low concentration of benzoic acid molecules in the bulk (Figure 2, panel a in the main text). Indeed, according to Ref. [8], the experimental solubility of BAH in water is rather low, 0.02 M. From another side, benzoic acid is a surface-active compound, which means that its concentration in the surface layer should exceed its concentration in the bulk layer. Indeed, in our experiments the signal of benzoic acid measured at the solution surface is approximately 30 times higher compare to bulk solution. This means, that the molecular density of benzoic acid in the surface layer should significantly exceed the molecular density of benzoic acid in the bulk in our simulations.

From another side, due to a finite size of our simulation box and finite simulation time we are unable to fully simulate the concentration profile in the system: most of benzoic acid molecules spend most of the simulation time in the surface layer. Moreover, due to the sampling problem, we cannot get a reliable statistics with a concentration of BAH around 0.01-0.02 M without a very large increase of the box size (that would lead to unreasonably large simulation costs). To overcome this difficulty we utilized different simulation technique, which allows us to obtain potential of mean forces (PMF) of a benzoic acid molecule travelling through a water slab. Generally speaking, the PMF represents the free energy profile along the z-direction. We performed these calculations because we would like to benchmark our force-field (the additional calculations can help us to understand whether our *model* benzoic acid is soluble or not).

We obtained the PMF as follows. Firstly, One neutral benzoic acid molecule was pulled through a slab of water (4 nm thick) with a constant rate of 10 nm/ns. System snapshotS (coordinates and velocities) were recorded over equal time intervals (0.02ns). Further, we allowed each of these system snapshots to equilibrate for 1 ns. The benzoic acid molecule was not allowed to move in z-direction. Forces acting in the system were recorded and averaged over the time. To reduce statistical noise we performed 5 independent simulations (replicas) with the same simulation protocol described above. To improve sampling, the benzoic acid molecule traveled left to right

in all odd replicas, and right to left in all even replicas. The initial configurations for each replica were generated as follows. We simulated an initial system configuration for 1000 steps at 500K. After the simulation was finished we assigned new velocities to all particles according to temperature 300K and minimized the energy. A unique random generator seed was used to assign velocities in each system.

Results of these simulations are represented on Figure S5 that shows the PMF (free energy profile) aligned with the water density profile. According to these calculations, the most energetically favorable position of benzoic acid in the water slab is the low-density surface layer of water. At the same time, this layer corresponds to the layer of the highest concentration of benzoic acid molecules represented at the figure 2, panel a in the main text. Therefore, we conclude that the results obtained by both simulation techniques coincide.

For an additional verification of our simulations force-field, we compare the experimental hydration free energy of BAH with the obtained value by the PMF calculations; that allows us to bridge our simulations and experimental data on benzoic acid solvation. Experimental hydration free energy of BAH is ΔG^{hyd}_{exp} = -15.6 kJ mol^{-1} [6], the hydration free energy value we obtained from our PMF calculations is -12.7 kJ mol^{-1}. That means that the hydration of our *model* benzoic acid is somewhat less favorable compare to the experimental values, which should give somewhat smaller solubility compared to the real benzoic acid. From another side, the difference is not very large and it is in the limit of typical experimental errors (that in the case of low soluble molecules, can exceed 4-5 kJ mol^{-1}, see e.g. discussion in [9]). We also would like to emphasize here that we did not aim to obtain a *quantitative* agreement with experiments, but rather to understand *qualitatively* the main phenomena associated with the interfacial behavior of benzoic acid at the vapor/water interface.

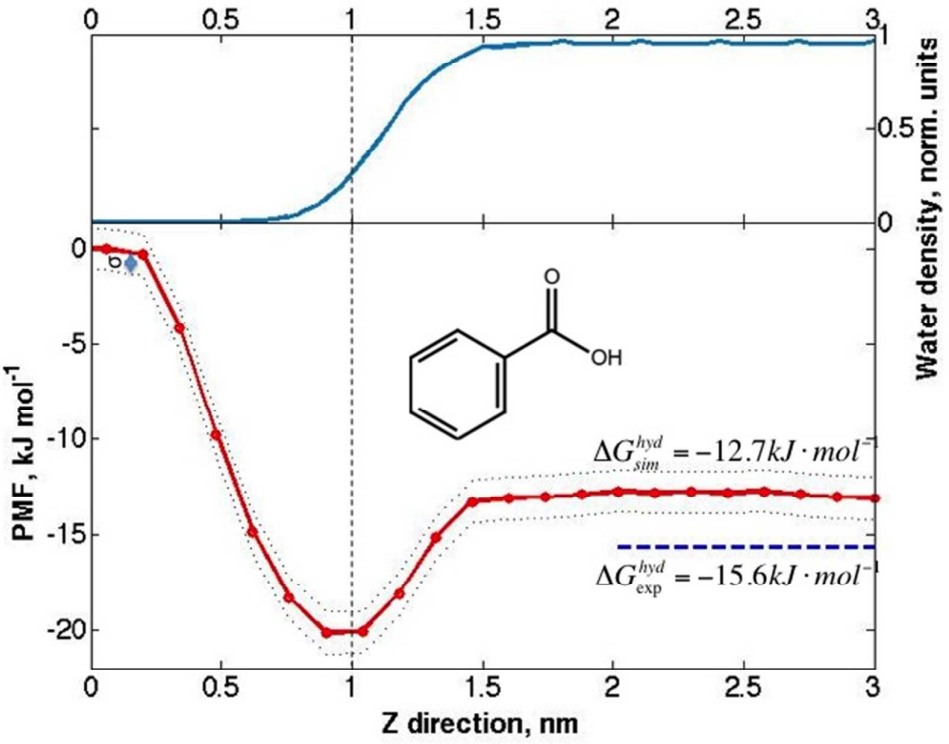

Figure S5. Water density profile (upper plot) and the PMF (free energy profile) of one BAH molecule travelling through the water slab (bottom plot).

Another important value that we can obtain from the free energy profile is the concentration ratios such as vacuum/surface, vacuum/bulk and surface/bulk [7]:

$$\frac{c_1}{c_2} = e^{-\frac{\Delta G_{12}}{RT}},$$

where c_1, c_2 are compound concentrations in phases 1 and 2; ΔG_{12} change of free energy upon transfer of a compound from phase 1 to phase 2; R – gas constant; T – temperature (300K). Results of our simulations are represented in the table S2. It is worth to highlight here that surface/bulk concentration ratio obtained from our simulation is in a good qualitative agreement with the experimental increase of benzoic acid peak intensity.

Table S2. Concentration ratios of benzoic acid in different phases (results from PMF calculations).

Phases	Concentration ratio
Surface/vapor	3500
Bulk/vapor	170
Surface/bulk	20

Comparison of our simulation results and experimental data allows us to conclude the following:

i) benzoic acid in our model is soluble in water at least at the low concentration limit;

ii) the model of benzoic acid represents well surface layer/bulk concentration ratios (at least *qualitatively*);

Coordinates of molecular structures used in this work

Structures for the molecular dynamics simulations were generated from scratch with Maestro 9.1 (Schrodinger suite 2010). The atom coordinates are represented below:

1.1) Benzoic acid

15

0MOLEC C4a 0 0.063 0.007 0.000
0MOLEC C4a 1 -0.008 0.130 0.000
0MOLEC C4a 2 -0.148 0.132 0.000
0MOLEC C4a 3 -0.221 0.011 0.000
0MOLEC C4a 4 -0.152 -0.112 0.000
0MOLEC C4a 5 -0.011 -0.114 0.000
0MOLEC H6a 6 0.047 0.223 0.000
0MOLEC H6a 7 -0.201 0.226 0.000
0MOLEC H6a 8 -0.329 0.013 0.000
0MOLEC H6a 9 -0.207 -0.204 0.000
0MOLEC H6a 10 0.040 -0.209 0.000
0MOLEC C2a 11 0.213 0.007 -0.000
0MOLEC O4a 12 0.270 -0.116 -0.000
0MOLEC O2a 13 0.280 0.110 -0.000
0MOLEC H8a 14 0.364 -0.105 -0.001
1.000 1.000 1.000

1.2) Benzoate

```
14
    0MOLEC  C4a    0   0.088  -0.001  -0.000
    0MOLEC  C4a    1   0.018   0.122   0.000
    0MOLEC  C4a    2  -0.123   0.123   0.000
    0MOLEC  C4a    3  -0.195   0.002   0.000
    0MOLEC  C4a    4  -0.125  -0.121   0.000
    0MOLEC  C4a    5   0.016  -0.122   0.000
    0MOLEC  H6a    6   0.074   0.214   0.000
    0MOLEC  H6a    7  -0.176   0.217   0.000
    0MOLEC  H6a    8  -0.303   0.003   0.000
    0MOLEC  H6a    9  -0.180  -0.214   0.000
    0MOLEC  H6a   10   0.070  -0.215  -0.000
    0MOLEC  C2a   11   0.239  -0.002  -0.000
    0MOLEC  O2a   12   0.298  -0.113  -0.000
    0MOLEC  O2a   13   0.300   0.107   0.000
   1.000   1.000   1.000
```

Afterwards, OPLS AA force field parameters were assigned with Desmond 2.4 [5]. They are represented in Table S1. The force field parameters were converted to Gromacs topology file.

SI references

[1] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *Molecular dynamics with coupling to an external bath* J. Chem. Phys. **81**, 3684 (1984)

[2] Ottosson N., Faubel M., Bradforth S. E., Jungwirth P. & Winter B., *Photoelectron spectroscopy of liquid water and aqueous solution: Electron effective attenuation lengths and emission-angle anisotropy*, J. Electron. Spectr. and Rel. Phen., **177**, 60 – 70 (2010)

[3] Faubel, M., *Photoelectron spectroscopy at liquid surfaces*, in Photoionization and Photodetachment, Part I, Ed. C. Y. Ng, World scientific, Singapore (2000)

[4] Söderström, J.; Mårtensson, N.; Travnikova, O.; Patanen, M.; Miron, C.; Sæthre, L. J.; Børve, K. J.; Rehr, J. J.; Kas, J. J.; Vila, F. D.; Thomas, T. D.; Svensson, S., *Non-stoichiometric intensities in core photoelectron spectroscopy*, Submitted to PRL (2012)

[5] Bowers, K. J., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., Shaw, D. E. *Scalable algorithms for molecular dymanics simulations on commodity clusters*, Proceedings of the ACM/IEEE Conference on Supercomputing, ACM Press, New York (2006)

[6] G. L. Perlovich, S. V. Kurkov, A. N. Kinchin, A. Bauer-Brandl, *Thermodynamics of solutions III: comparison of the solvation of (+)-naproxen with other NSAIDs,* European Journal of Pharmaceutics and Biopharmaceutics **57** 411–420 (2004)

[7] R. Vácha , P. Jungwirth , J. Chen, K. Valsaraj, *Adsorption of polycyclic aromatic hydrocarbons at the air–water interface: Molecular dynamics simulations and experimental atmospheric observations*, Phys. Chem. Chem. Phys., **8**, 4461–4467 (2006)

[8] taken from GESTIS Substance Database (http://gestis-en.itrust.de/nxt/gateway.dll/gestis_en/022810.xml?f=templates$fn=default.htm$3.0)

[9] E.L. Ratkova and M.V. Fedorov (2011). Combination of RISM and Cheminformatics for Efficient Predictions of Hydration Free Energy of Polyfragment Molecules: Application to a Set of Organic Pollutants. // *Journal of Chemical Theory and Computation,* **7** (5), 1450