Supplementary Information

Sensing the Acidity of Hydrogen Bond Networks

Tirthick Majumder[†1], R. Allen LaCour[†2,3], Archishman Sarkar[1], Teresa Head-Gordon[2,3*], Jahan M. Dawlaty[1*]

[1]Department of Chemistry, University of Southern California, Los Angeles, CA 90089, USA

[2]Kenneth S. Pitzer Theory Center and Department of Chemistry, University of California, Berkeley, CA, 94720 USA

[3]Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720 USA

[†]These authors contributed equally to this work.

Methods:

Experimental: The concentration of MeCN was 500 mM in the Raman measurements. We collected each Raman spectrum from within a glass vial containing the sample, using the HORIBA XploRA micro-Raman microscope with a 532 nm excitation laser.

The concentration of aniline in our NMR measurements was 110 mM. We used D_6-DMSO in the mixtures instead of DMSO for the NMR measurements. All spectra were recorded with a Varian 500 MHz spectrometer.

We prepared 4-cyanobenzo thiolate (4-CBT) by using NaOMe to deprotonate 100 mM 4-mercapto benzonitrile (4-MBN). We collected Raman spectra of the titration samples using the same setup as mentioned before. The spectra were interpolated using splines and then fitted to 2 to 3 Gaussians as applicable using MATLAB.

Computational: We generated the data for training our MLIP using *ab initio* molecular dynamics simulations. We used the Quickstep module[1] of the CP2K software package[2] to perform the *ab initio* molecular dynamics simulations. To obtain a robust MLIP,[3] we generated configurations under multiple conditions. First, we generated configurations using the revPBE functional with Grimme's D3 correction for dispersion.[4] Compared to similar levels of theory, revPBE has been shown to work well with water.[5] We performed the simulations using a Gaussian plane wave basis set using the TZV2P basis and plane waves with a cutoff of 400 Ry. We used the Goedecker-Teter-Hutter pseudopotentials for revPBE to represent the core electrons.[6] We optimized the orbitals using the DIIS method[7] with a convergence criteria of 10^{-6}. For DMSO-water mixtures, we ran NVT simulations by linearly interpolating experimentally reported densities.[8] For DMSO-HFIP mixtures, we ran NVT simulations at densities measured by us (Table S1). Each DMSO-water AIMD simulation was run for at least 100 ps and with three replicates; each of the larger DMSO-HFIP AIMD simulations was run for at least 60 ps. The temperature was set to 300 K. Second, we ran AIMD simulations of only monomers using the same level of theory at the higher temperature of 3000 K to obtain highly distorted geometries. Third, we ran four simulations using the extended tight binding method [9](xTB) at the higher temperature of 2000 K. These were at DMSO to cosolvent ratios of 6:48 and 48:6 for both water and HFIP mixtures.

It is important for MLIPs to be trained on data computed at the same level of theory. Thus, for all three cases listed above, we then recomputed the forces and energies using the same schema as our first AIMD simulations except with the plane wave cutoff raised to 600 Ry. This approach, in which we compute the forces and energies at a higher level of theory from configurations generated from simulations at a lower level of theory, is an effective way to generate a diverse set of molecular configurations. We do not expect the configurations generated with xTB to lead to unphysical biases in our dataset because, if the configurations are high energy, our MLIP will learn to avoid them.

Our MLIPs were trained using the neural network potential Newtonnet[10]. We trained three separate models for both the DMSO-water and DMSO-HFIP systems. We typically obtained mean absolute force and energy errors on the order of 35 to 50 meV/Å and 0.0005 meV per atom, respectively, on the test set.

The simulations shown in Figure 2 were run using the MLIPs. Specifically, we used an ensemble of all three models for each mixture to run the MD simulations. We used 53 solvent molecules and one acetonitrile molecule in each simulation. We equilibrated each system for at least 50 ps before conducting a production run of 1 ns. Each system was simulated using a

Langevin thermostat at 300 K with a time constant of 10 ps. We simulated four replicas at each solvent ratio. We used the atomic simulation environment[11] (ASE) to run the simulations.

To generate the spectra shown in Figures 3a and b, we dumped the velocities of the N and C atoms in the MeCN molecule every 2 fs. To obtain a finer temporal resolution, we then used Akima interpolation[12] to obtain the velocity at 1/15 fs intervals. We found interpolating to smaller time intervals was equivalent to dumping the velocities more frequently but resulted in significantly less storage. We then took the Fourier transform of the autocorrelation of the velocities. We used a 6 ps Gaussian window function to smooth the autocorrelation.

We originally found a superharmonic to be an issue in all computationally produced spectra. We found that changing the mass of the H could alleviate the issue. Changing the mass of the H to deuterium resulted in the CH stretching mode overlapping with the CN region, whereas changing the mass to that of tritium alleviated the issue in all spectra while minimally affecting the CN stretch. See Figure S5 for a comparison. For this reason, all production simulations using the MLIPs (which include all computational results shown in the main text) are performed with tritiated MeCN. We note that none of the AIMD or xTB simulation used to generate data for training the MLIPs used tritiated MeCN, but this is not an issue because atomic masses do not influence the distribution of configurations sampled under classical dynamics or the forces or energies of those configurations. We also further validate and analyze our systems in the radial distribution functions shown in Figures S6-S8.

Hybrid functionals were not used because our AIMD simulations with DMSO and HFIP frequently exceeded 500 atoms, which is prohibitively slow for a hybrid functional. We note that, at our level of DFT, revPBE with the D3 dispersion correction is generally the recommended functional.[13] Our work further supports its use considering that we could capture the trend in acetonitrile frequency for the systems examined here.

We used the Freud[14] and Signac[15] software packages to analyze the trajectories and organize data, respectively.

1. **MeCN as a Hydrogen Bonding Probe:**

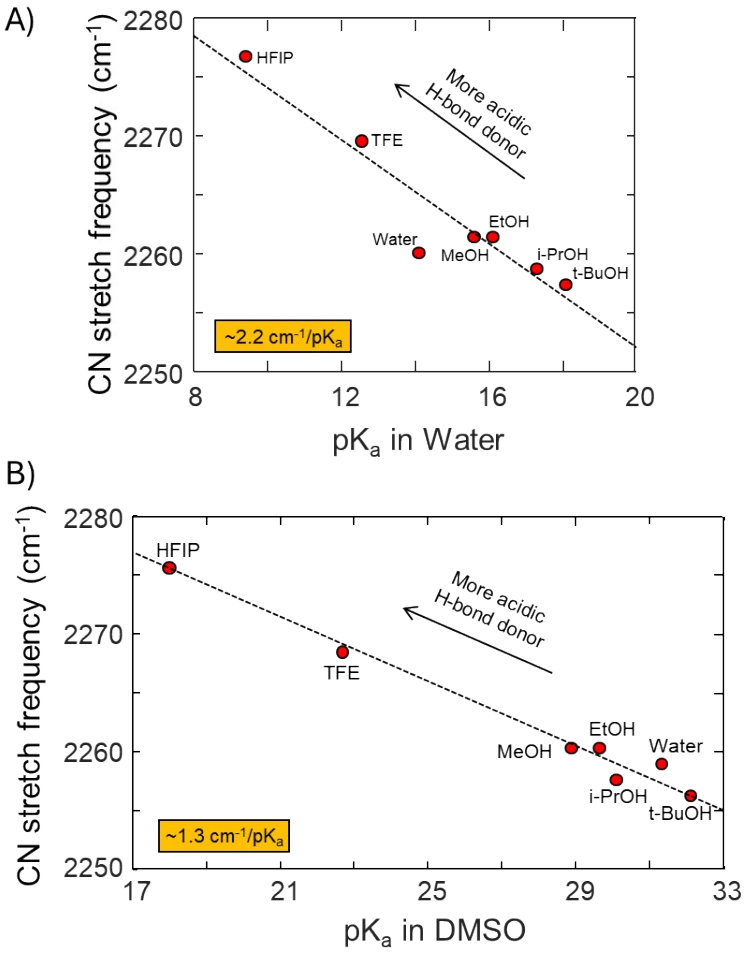


Figure S1: Relation between frequency shift of MeCN and pK_a of solvent in protic solvents containing the OH functional group. Our results resemble previous reports.[16]

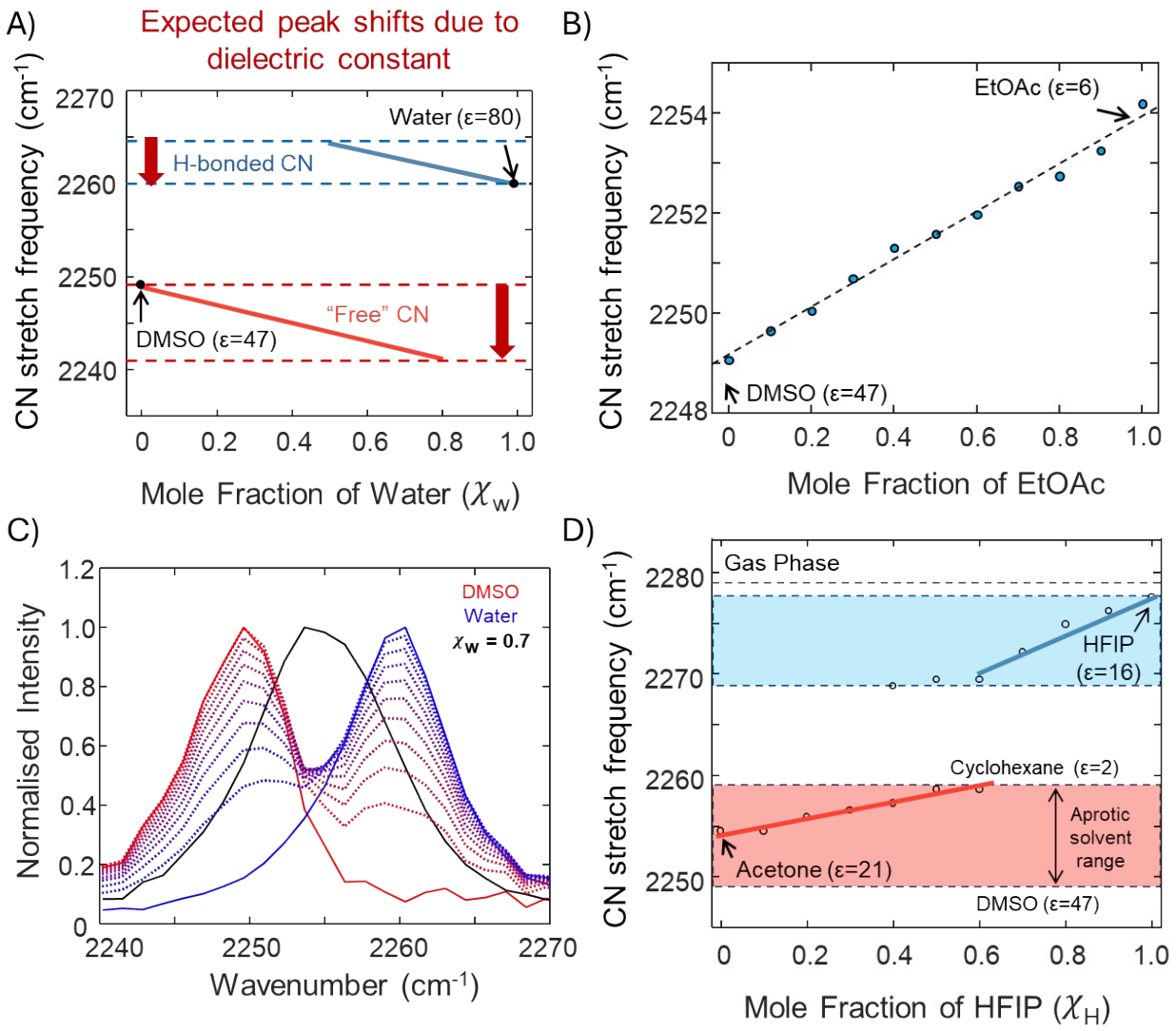


Figure S2: (A) Expected frequency shift considering only dielectric effects and ignoring any 'network' effect. B) Shift of the CN stretch frequency of MeCN as a function of the mole fraction of EtOAc in DMSO-EtOAc mixtures. C) Attempted reconstruction of MeCN spectrum in 70% water, from the spectra in pure DMSO and pure water. All linear combinations of the spectra in pure DMSO and pure water fail to reproduce the spectrum of the mixture D) Peak frequency of the CN stretch of MeCN as a function of the mole fraction of HFIP in DMSO-HFIP mixtures

2. Analysis of MeCN Frequency Shifts in DMSO-water mixtures:

The CN stretch frequency of acetonitrile changes with the mole fraction of water (χ_w) as shown in figure 1B. Note that the dielectric constant of DMSO-water mixtures increases continuously with χ_w.(Ref 87 of main text) Therefore, a purely dielectric effect would result in a continuous red shift with χ_w as indicated in Figure S3A. We verified this experimentally using an aprotic-aprotic mixture of DMSO and EtOAc and observed a continuous linear blue shift from pure DMSO to pure EtOAc.(Figure S2). This implies that the continuous blue shift observed in our experiments is due to a non-dielectric effect (or specific interaction), which in this case, is hydrogen bonding. Also note that the linewidths were broader in the intermediate mole fractions (χ_w = 0.7, 0.8, 0.9) as opposed to terminal mole fractions (Figure S2), which suggests the presence of two subpopulations, namely, the hydrogen-bonded and non-hydrogen-bonded or 'free' MeCN. This inference is further confirmed by the presence of two MeCN peaks in DMSO-HFIP mixtures (Figure 1C). Additionally, changes in the hydrogen bonding strength of water networks upon dilution with DMSO is expected to be reflected in the stretching frequency of the CN hydrogen bonded to that water network. To summarize, the peak frequency shift is a cumulative effect of the following factors: i) existence of two populations corresponding to hydrogen-

bonded and 'free' MeCN molecules ii) dielectric or solvent reaction field effect on both these populations iii) changes in hydrogen bonding strength with growth of the HBN ("network effect").

Keeping all this in mind, we proceed to analyze the χ_w dependence of frequency shifts. At $\chi_w = 0$, MeCN is solvated entirely by DMSO molecules, and its frequency is 2249 cm^{-1}. As we gradually increase the water content, some of the MeCN molecules form hydrogen bonds to water clusters. Eventually, it reaches a point where all the MeCN molecules are hydrogen bonded and continues like that till $\chi_w = 1$, that is, pure water. Therefore, as we traverse χ_w from 0 to 1, the hydrogen-bonded MeCN population is initially 0 up to a certain threshold and then increases continuously until it reaches a high χ_w where all the MeCN molecules are hydrogen bonded. Since hydrogen bonding causes a blue shift, an increase in the hydrogen-bonded population is expected to cause a blue shift of the average frequency (comprising both hydrogen-bonded and non-hydrogen-bonded populations). Here, one may be tempted to attribute the continuous non-linear blue shift of the peak frequency (Figure 2B) entirely to the growth of the hydrogen-bonded population, without invoking the "network effect" of the HBN at all. In the following arguments, we will demonstrate why this hypothesis is inadequate to describe the spectra of Figure 2A. First, let us ignore any dielectric and "network" effects. Then, the MeCN spectrum in a DMSO-water mixture of any arbitrary χ_w should be expressible as a linear combination of the spectra in pure DMSO and pure water. We performed such linear combinations as an exercise and showed that the spectrum of $\chi_w = 0.7$ cannot be reproduced by this method. (Figure S2C) This is because the peak frequencies in pure DMSO and pure water are significantly separated (~11 cm^{-1}), so that their linear combinations usually produce two peaks (Figure S2C), contrary to a single peak observed in the DMSO-water mixtures (Figure 2A). This implies that, along with the increase in the number of hydrogen bonded MeCN molecules, the average frequency of all the hydrogen-bonded MeCN molecules must also shift to depict the spectra of Figure 2A. At this stage, it remains to be confirmed whether this shift is borne out of the "network effect" or can be explained by mere dielectric effects. Note that, one may attempt to fit the spectra to two Gaussians, corresponding to the hydrogen-bonded and "free" MeCN molecules. However, this approach will be erroneous as the MeCN spectra even in pure DMSO and pure water are asymmetric and therefore, incorrectly represented by a Gaussian. Adding an arbitrary number of Gaussians would undeniably improve the fit but due to lack of proper assignment of each Gaussian to a subpopulation, we refrained from this approach. Instead, we estimated the proportions of hydrogen-bonded and "free" MeCN in the mixtures from the partitioning coefficient of MeCN between DMSO and water. (Refer to SI Section 2).

Continuing our line of reasoning, let us now include the dielectric effect and only ignore the "network effect". The dielectric constant of DMSO-water mixtures continuously increases with χ_w, albeit non-linearly. (Ref) Therefore, both the "free" and hydrogen-bonded MeCN populations are expected to red shift with increasing χ_w. The overall χ_w dependence of the peak frequency would be as follows (FigureS2A): It would begin at 2249 cm^{-1} at $\chi_w = 0$ (pure DMSO), then slowly red shift with χ_w up until there is a detectable hydrogen-bonded population. Note that the hydrogen-bonded population is also expected to red shift due to increase in dielectric constant up until $\chi_w = 1$ (pure water), which is at 2260 cm^{-1}. Therefore, at its onset (let us call it χ_{HB}), the HB population should be blue shifted from 2260 cm^{-1}. Furthermore, the peak-to-peak separation between the HB and the "free" MeCN populations at χ_{HB} would be larger than the separation between the pure water and pure DMSO spectra. Their linear combination would result in two peaks, in accordance with the discussion of Figure S2C. The expected behavior of the peak frequency in the absence of the "network effect" is depicted in Figure S3, which looks nothing like the experimental data (Figure 2B). These arguments illustrate the fact that only population exchange between HB and "free" MeCN, along with changes in dielectric with χ_w are inadequate to explain the spectra of Figure 2A. In fact, these are likely minor contributors to the shift because the expected direction of shift (Figure S2A) is opposite to the data (Figure 1B).

Therefore, the only possibility that remains is the "network effect". Growth of the HBN increases the hydrogen bond donation strength which further blue shifts the CN stretch of the hydrogen bonded subpopulation. This results in a continuous overall blue shift of the CN stretch peak (Figure 1B).

3. Analysis of MeCN Frequency Shifts in DMSO-HFIP mixtures:

The pK_a's of HFIP in DMSO and in water are 17.9[17] and 9.3, respectively. These numbers correspond to the dilute limit of HFIP in the corresponding solvents. Neat HFIP, which is likely to have a lower pK_athan its solution in DMSO or water, is often used to promote acid catalyzed reactions.[18] It has been suggested that dimers and trimers are involved in many HFIP-mediated

reactions.[19] Understanding the changes in H-bond strength and its effect on the $^{p}K_{a}$ of HFIP upon systematic dilution with aprotic and protic solvents will shed more light on its catalytic activity.

Hydrogen bonded and "free" MeCN subpopulations exist as two discernable CN stretch peaks (Figure 1C) both of which blue shift with increasing χ_H (Figure 1D). Naively, the blue shift of both peaks could be attributed to the decrease in dielectric constant from DMSO to HFIP, without invoking any H-bond network effects. To address this issue, we performed a control experiment using Acetone instead of DMSO. Despite similar dielectric constants of acetone and HFIP, the H-bonded peak blue shifts significantly in acetone-HFIP mixtures (Figure S2 D). This confirms that the shift of the H-bonded peak is due to the growth of the H-bonded network and not merely a dielectric effect.

4. Analysis of ^{1}H NMR chemical shifts of aniline:

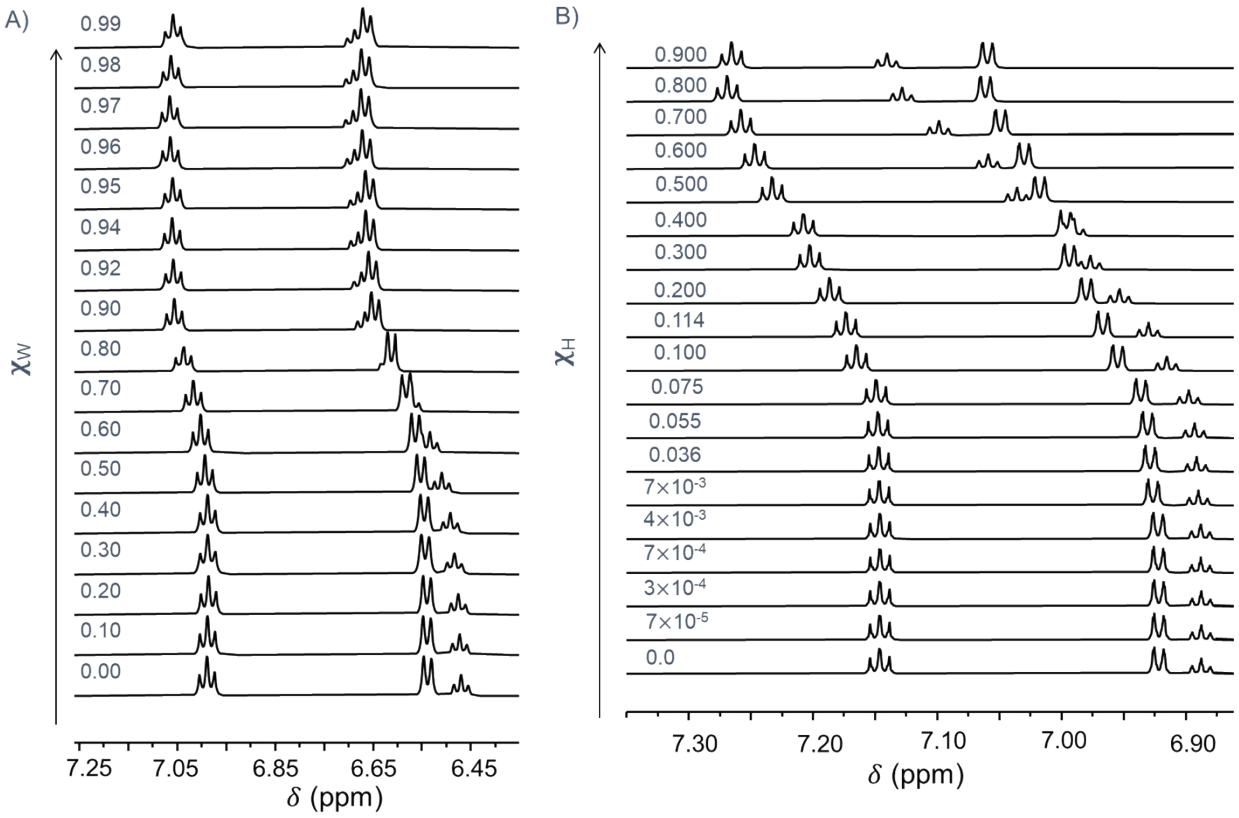


Figure S3: NMR titrations curves showing the ring protons of aniline when titrated against (A) Water and (B) HFIP in a D_6-DMSO background.

The NMR titration spectra are shown in Figure S3. The fraction of hydrogen bonded anilines, f, at any mole fraction of water or HFIP is calculated using the following equation

$$f = \frac{\delta - a}{b - a},$$

where, δ is the chemical shift of the ring proton at the desired mole fraction of water or HFIP, b and a are the chemical shifts of the ring proton for fully H-bonded aniline and 'free' aniline respectively. The equilibrium for hydrogen bonding is

$$PhNH_2 + ROH \longrightarrow PhNH_2 \cdots HOR$$

$$K_{eq} = \frac{[PhNH_2 \cdots HOR]}{[PhNH_2][ROH]}$$
$$= \frac{c}{(c_1 - c)(c_2 - c)},$$

where, c is the concentration of H-bonded aniline, c_1 is the total concentration of aniline (both H-bonded and free), and c_2 is the total concentration of ROH (water or HFIP). We know that

$$\frac{c}{c_1} = f = \frac{\delta - a}{b - a}$$
$$\Rightarrow c = c_1 \frac{\delta - a}{b - a}$$

Using the above expression of c we calculated the K_{eq}.

The dependence on ROH concentration implies that our reference state is 1M ROH instead of pure ROH. We choose 1M ROH because our ROH concentration varies across the solvent mixtures. The errors in K_{eq} are calculated from the measurement errors in concentrations.

The equilibrium constant for hydrogen bonding is then converted to free energy using the relation: $\Delta G_{H.B.} = -RTlnK_{eq}$

In the dilute limit, the free energy of hydrogen bonding is defined to be $\Delta G^{(0)}$. However, because in our work, we explicitly scan over a large range of concentrations, from dilute all the way to pure solvents, therefore, we use the notation $\Delta G_{H.B.}$. Note that $\Delta G_{H.B.}$ subsumes within it the excess energy due to the network effect and correspondingly the equilibrium constant K_{eq}. An alternative common choice is to keep the free energy as $\Delta G^{(0)}$, but modify the concentrations within the equilibrium constant by activity coefficients.

5. Partition Coefficients of MeCN:

We estimated the partitioning of MeCN into hydrogen bonded and "free" subpopulations in DMSO-water mixtures by measuring its partition coefficient between DMSO and water. Since DMSO and water are miscible, we measured the partition coefficient between water and hexane, and ratioed it to the partition coefficient between DMSO and hexane. For each measurement, we first prepared an MeCN solution (in 1 mL water or DMSO). Then, we carefully added 1mL hexane to the solution to minimize any mixing and left it unperturbed for 4 days. We separately pipetted out the hexane and aqueous (or DMSO) phases and collected their Raman spectra. We related the area under the CN stretch to the population of MeCN, converted solvent volume to number of moles and obtained that the partitioning from hexane to DMSO is ~ 3 times that from hexane to water for the same number of moles.

6. Analysis of acid-base titration:

The Raman cross-sections of the deprotonated and neutral forms of 4-MBN were independently measured under identical experimental conditions. The ratio of the Raman cross section of the deprotonated to the neutral form was ~1.86 which was used in the calculation of the equilibrium constant. For each of the spectra shown in figure 4, the concentrations of the deprotonated ($CNPhS^-$) and neutral 4-MBN ($CNPhSH$) were calculated by normalizing the area under the Gaussian fits to their respective cross-sections. The equilibrium constant was calculated as follows:

$$CNPhS^- + ROH \longrightarrow CNPhSH + RO^-$$
$$K_{eq} = \frac{[CNPhSH][RO^-]}{[CNPhS^-][ROH]}$$
$$= \frac{c^2}{(c_1 - c)(c_2 - c)},$$

where, c is the concentration of $CNPhSH$ or RO^-, c_1 is the initial concentration of $CNPhS^-$, and c_2 is the initial concentration of ROH (HFIP). The ratio of protonated to unprotonated base, $\frac{c}{c_1 - c}$ is obtained from the Gaussian fits, which is then solved for c and plugged into the above equation for K_{eq}. Similar to the equilibrium constant of hydrogen bonding, the $K_{eq}$for acid-base titration is also referenced to 1M HFIP. The errors in K_{eq} are calculated by propagating the fitting errors.

7. Simulations:

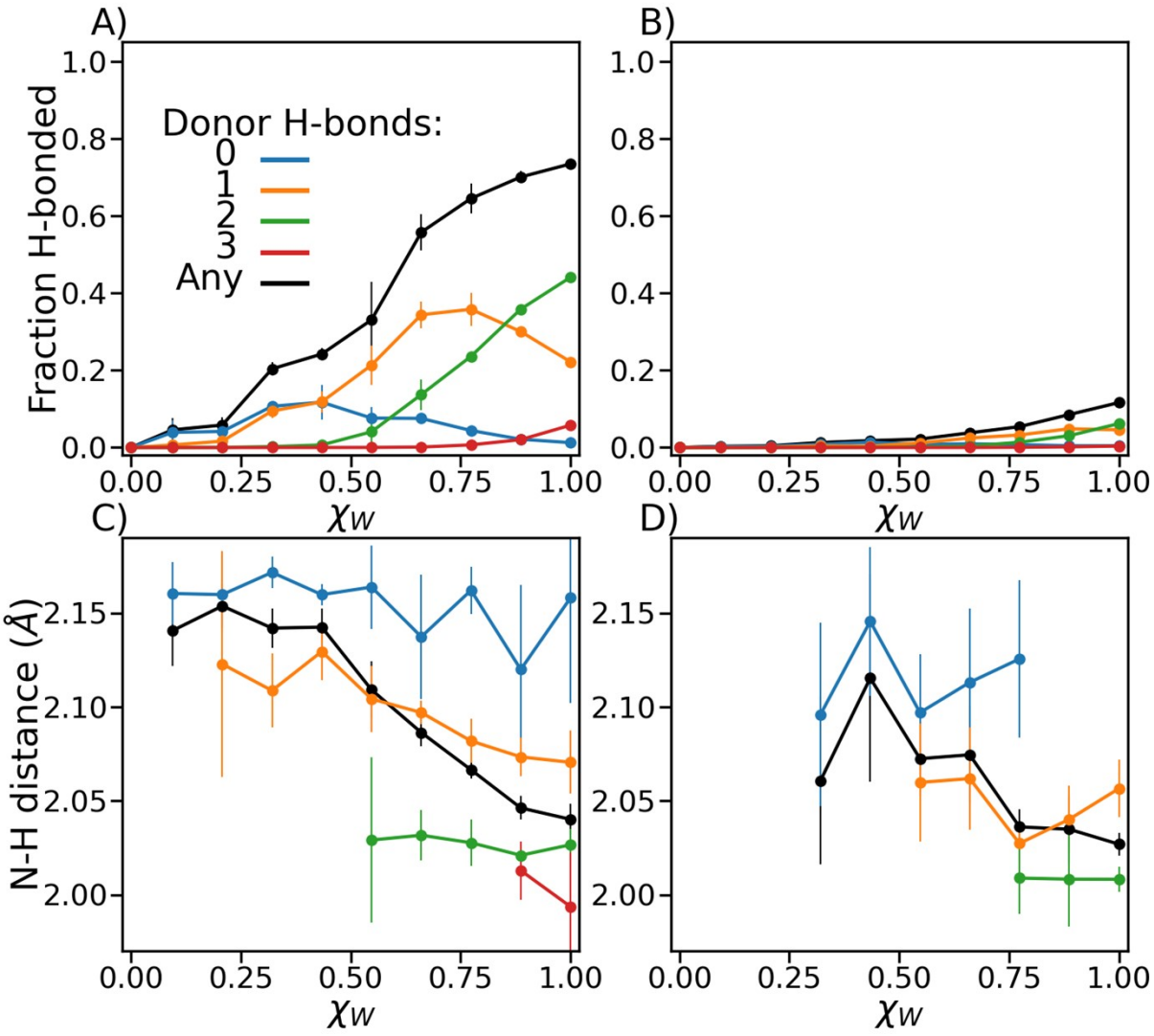


Figure S4: H-bond strength of water. Here we further analyzed the results for water presented in Figure 2.D, F). Specifically, we divided the results by whether the water donating an H-bond to MeCN was also donating to another molecule (A,C) or not (B,D). The fractions shown in panels A and B sum to those shown in Figure 2.D. The fractions are much greater in panel A, showing that the water donating to MeCN is almost always donating an H-bond to another molecule. The N-H distances shown in panel C are very similar to those shown in Figure 2F. The N-H distances shown in panel D are also similar but have more noise due to fewer configurations being included in the average. In panel C we follow the main text and only show points where the MeCN spends over 1% of its time in that environment. Because so few configurations are measured in panel D, we show all points where MeCN spends over 0.5 % of its time.

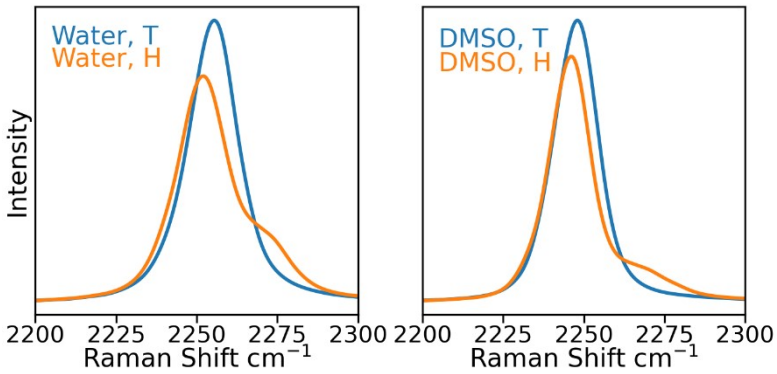
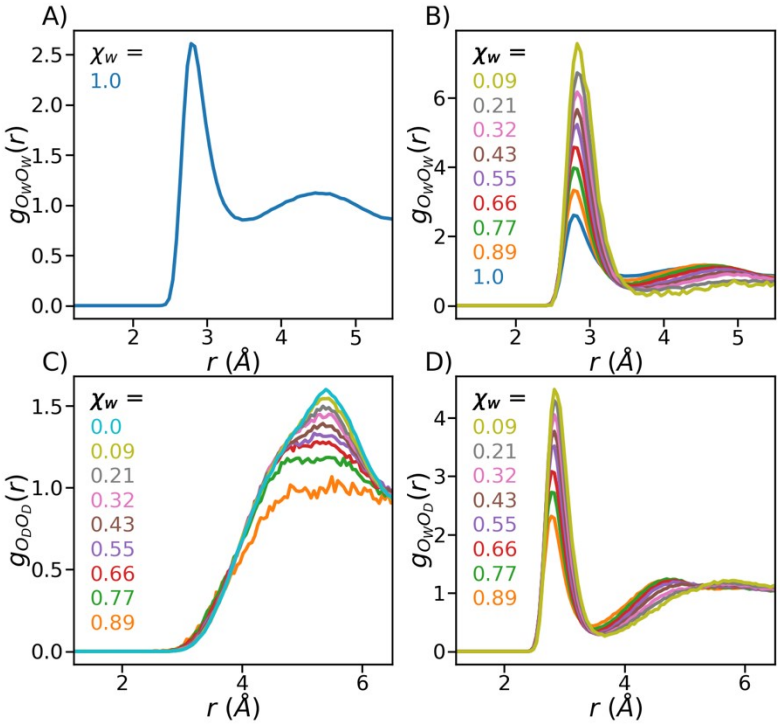


Figure S5: Tritiating MeCN in simulation. When using MeCN with H (the "Water, H" and "DMSO, H" spectra) in simulation, we observed a small shoulder in the spectra. We attribute this to a superharmonic vibration involving the CH stretches of the MeCN. The shoulder hampered the interpretability of the spectra. Substituting D for H did not solve the issue because then the CDstretch region overlapped with the CN stretch. Substituting T for the H eliminated the issue (the "Water, T" and "DMSO, T" spectra)



without substantially perturbing the C-N stretch. We only do the substitution for H in MeCN; all other H are not changed.

Figure S6: Radial distribution functions ($g(r)$) in water/DMSO systems at different stoichiometries. (A) The $g(r)$ between the oxygen of water (O_W) when the solvent is pure water. There is still one acetonitrile in the system, but the $g(r)$ resembles those reported previously.[20] (B) The same $g(r)$ except at different solvent compositions. (C) The $g(r)$ between the oxygens of DMSO (O_D). (D) The $g(r)$ between the oxygen of water and those of DMSO. The rise of the first peak in $g_{O_WO_W}(r)$ and $g_{O_WO_D}(r)$ with increasing DMSO concentration has been reported previously and attributed to excluded volume effects.[21]

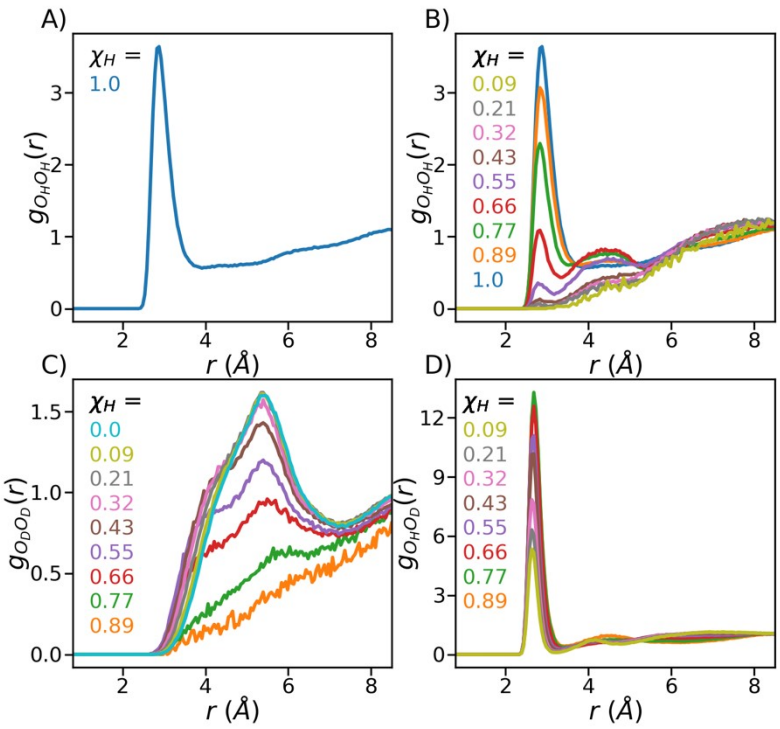


Figure S7: Radial distribution functions ($g(r)$) in HFIP/DMSO systems at different stoichiometries. The larger size of these systems allowed us to use a longer cutoff in the $g(r)$. (A) The $g(r)$ between the oxygens of HFIP (^{O_H}) when the solvent is pure HFIP. There is still one acetonitrile in the system. (B) The same $g(r)$ except at different solvent compositions. (C) The $g(r)$ between the oxygen of DMSO (^{O_D}). (D) The $g(r)$ between the oxygen of water and those of DMSO. Unlike with water, the first peak in the $g_{O_HO_H}(r)$ and $g_{O_HO_D}(r)$ decline with increasing DMSO concentration. We attribute this to that fact that HFIP has a hydrophobic group and can only donate one H-bond.

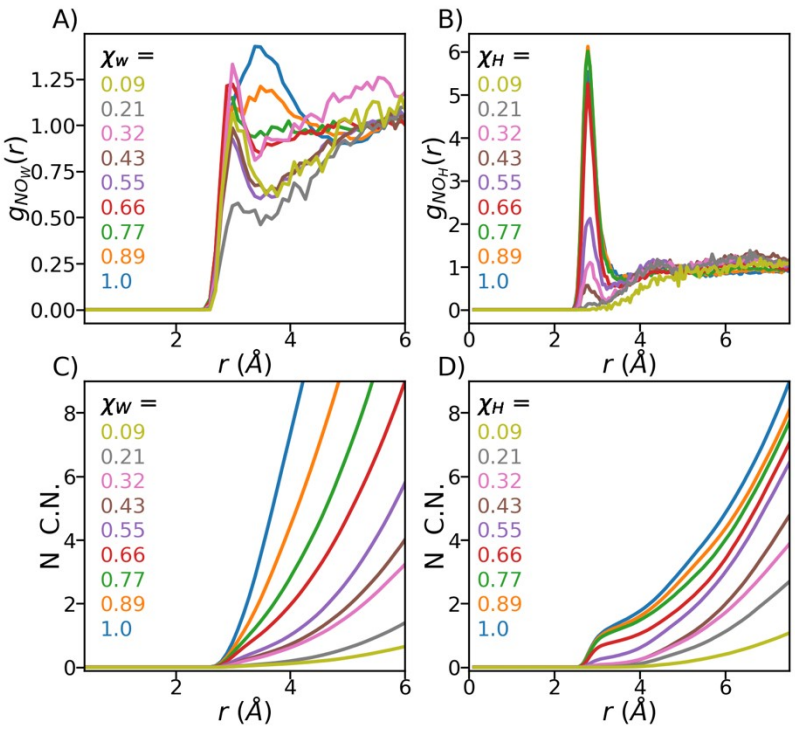


Figure S8: Radial distribution functions between MeCN and the H-bonding species. (A) The $g(r)$ between the oxygen of water and the nitrogen (N) of the MeCN in water/DMSO systems. (B) The same except for the oxygen of HFIP in HFIP/DMSO systems. (C) The average coordination number (C.N.) of the oxygen of water around the nitrogen of the MeCN. (D) The same except for the oxygen of HFIP. The behavior for HFIP in panels (B,D) is straightforward, with the probability of HFIP coordinating to the N increasing with increasing HFIP concentration. The behavior for water in panels (A,C) is more complex. At low at water concentration, the MeCN is likely to interact with a single isolated water, yielding a single narrow peak at the H-bond distance. At higher water concentrations, it is more likely to H-bond a water network. This yields a broader peak that is slightly shifted to longer distances. However, as shown in Figure 3F in the main text and the C.N. plot shown in panel C, the average distance to the nearest water decreases with increasing concentration.

8. Additional:

Mole Fraction DMSO	Density (g/mL)
0.0	1.5325
0.05	1.417
0.1	1.4
0.15	1.4425
0.2	1.4555
0.3	1.432

0.4	1.406
0.5	1.3565
0.6	1.3095
0.7	1.2735
0.8	1.2055
0.9	1.1445
1.0	1.1105

Table S1. Density of DMSO/HFIP Mixtures. We prepared the mixture of desired mole fraction and weighed 1 mL of it to obtain the density. The values shown are averaged over three measurements.

References:

(1) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Commun.* **2005**, *167* (2), 103–128. https://doi.org/10.1016/j.cpc.2004.12.014.

(2) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. Cp2k: Atomistic Simulations of Condensed Matter Systems. *WIREs Comput. Mol. Sci.* **2014**, *4* (1), 15–25. https://doi.org/10.1002/wcms.1159.

(3) Yuan, E. C.-Y.; Head-Gordon, T. Teachers That Teach the Irrelevant: Pre-Training Machine Learned Interaction Potentials with Classical Force Fields for Robust Molecular Dynamics Simulations. arXiv April 8, 2026. https://doi.org/10.48550/arXiv.2509.14205.

(4) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104. https://doi.org/10.1063/1.3382344.

(5) Ohto, T.; Dodia, M.; Xu, J.; Imoto, S.; Tang, F.; Zysk, F.; Kühne, T. D.; Shigeta, Y.; Bonn, M.; Wu, X.; Nagata, Y. Accessing the Accuracy of Density Functional Theory through Structure and Dynamics of the Water–Air Interface. *J. Phys. Chem. Lett.* **2019**, *10* (17), 4914–4919. https://doi.org/10.1021/acs.jpclett.9b01983.

(6) Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B* **1996**, *54* (3), 1703–1710. https://doi.org/10.1103/PhysRevB.54.1703.

(7) Weber, V.; VandeVondele, J.; Hutter, J.; Niklasson, A. M. N. Direct Energy Functional Minimization under Orthogonality Constraints. *J. Chem. Phys.* **2008**, *128* (8), 084113. https://doi.org/10.1063/1.2841077.

(8) LeBel, R. G.; Goring, D. A. I. Density, Viscosity, Refractive Index, and Hygroscopicity of Mixtures of Water and Dimethyl Sulfoxide. *J. Chem. Eng. Data* **1962**, *7* (1), 100–101. https://doi.org/10.1021/je60012a032.

(9) Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All Spd-Block Elements ($Z = 1–86$). *J. Chem. Theory Comput.* **2017**, *13* (5), 1989–2009. https://doi.org/10.1021/acs.jctc.7b00118.

(10) Haghighatlari, M.; Li, J.; Guan, X.; Zhang, O.; Das, A.; Stein, C. J.; Heidar-Zadeh, F.; Liu, M.; Head-Gordon, M.; Bertels, L.; Hao, H.; Leven, I.; Head-Gordon, T. NewtonNet: A

Newtonian Message Passing Network for Deep Learning of Interatomic Potentials and Forces. *Digit. Discov.* **2022**, *1* (3), 333–343. https://doi.org/10.1039/D2DD00008C.
(11) Hjorth Larsen, A.; Jørgen Mortensen, J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Bjerre Jensen, P.; Kermode, J.; Kitchin, J. R.; Leonhard Kolsbjerg, E.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Bergmann Maronsson, J.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schütt, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. W. The Atomic Simulation Environment—a Python Library for Working with Atoms. *J. Phys. Condens. Matter* **2017**, *29* (27), 273002. https://doi.org/10.1088/1361-648X/aa680e.
(12) Akima, H. A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures. *J ACM* **1970**, *17* (4), 589–602. https://doi.org/10.1145/321607.321609.
(13) Ohto, T.; Dodia, M.; Xu, J.; Imoto, S.; Tang, F.; Zysk, F.; Kühne, T. D.; Shigeta, Y.; Bonn, M.; Wu, X.; Nagata, Y. Accessing the Accuracy of Density Functional Theory through Structure and Dynamics of the Water–Air Interface. *J. Phys. Chem. Lett.* **2019**, *10* (17), 4914–4919. https://doi.org/10.1021/acs.jpclett.9b01983.
(14) Ramasubramani, V.; Dice, B. D.; Harper, E. S.; Spellings, M. P.; Anderson, J. A.; Glotzer, S. C. Freud: A Software Suite for High Throughput Analysis of Particle Simulation Data. *Comput. Phys. Commun.* **2020**, *254*, 107275. https://doi.org/10.1016/j.cpc.2020.107275.
(15) Adorf, C. S.; Dodd, P. M.; Ramasubramani, V.; Glotzer, S. C. Simple Data and Workflow Management with the Signac Framework. *Comput. Mater. Sci.* **2018**, *146*, 220–229. https://doi.org/10.1016/j.commatsci.2018.01.035.
(16) Xu, J.; Deng, G.; Wang, Y.-T.; Guo, H.-Y.; Kalhor, P.; Yu, Z.-W. Local Acid Strength of Solutions and Its Quantitative Evaluation Using Excess Infrared Nitrile Probes. *J. Phys. Chem. Lett.* **2020**, *11* (3), 1007–1012. https://doi.org/10.1021/acs.jpclett.9b03804.
(17) Pka-Compilation-Reich-Bordwell.Pdf. https://organicchemistrydata.org/hansreich/resources/pka/pka_data/pka-compilation-reich-bordwell.pdf (accessed 2024-05-17).
(18) *"Haliranium Ion"-Induced Intermolecular Friedel-Crafts Alkylation in HFIP: Synthesis of β,β-Diaryl α-Halo carbonyl Compounds*. https://doi.org/10.1002/ejoc.202100823.
(19) Berkessel, A.; Adrio, J. A.; Hüttenhain, D.; Neudörfl, J. M. Unveiling the "Booster Effect" of Fluorinated Alcohol Solvents: Aggregation-Induced Conformational Changes and Cooperatively Enhanced H-Bonding. *J. Am. Chem. Soc.* **2006**, *128* (26), 8421–8426. https://doi.org/10.1021/ja0545463.
(20) Pestana, L. R.; Mardirossian, N.; Head-Gordon, M.; Head-Gordon, T. Ab Initio Molecular Dynamics Simulations of Liquid Water Using High Quality Meta-GGA Functionals. *Chem. Sci.* **2017**, *8* (5), 3554–3565. https://doi.org/10.1039/C6SC04711D.
(21) Zhang, N.; Li, W.; Chen, C.; Zuo, J. Molecular Dynamics Simulation of Aggregation in Dimethyl Sulfoxide–Water Binary Mixture. *Comput. Theor. Chem.* **2013**, *1017*, 126–135. https://doi.org/10.1016/j.comptc.2013.05.018.