Supporting Information

for

Humidity effect on the first steps of atmospheric new particles formation: computational study of hydrated molecular clusters

Ivo Neefjes, Yosef Knattrup, Haide Wu, Georg Baadsgaard Trolle, Jonas Elm, and Jakub Kubečka*

Aarhus University, Department of Chemistry, Langelandsgade 140, Aarhus, 8000, Denmark

E-mail: ja-kub-ecka@chem.au.dk

Phone: +420 724946622

S1 Configurational Sampling Procedure

The geometries of molecular clusters were found through configurational sampling exploration using the JK framework.[1] Here, we further discuss each dataset.

Microhydrated monomer and dimer clusters:

For each of the (acid and/or base)$_{0-2}$W$_{0-5}$ clusters, the following steps were performed:

- In three independent simulations, 300 guess structures were generated (population) with the rigid-molecule optimizer in ABCluster.[2,3] This structure set was optimized over 100 steps (generations), and the lowest 300 structures (local minima) were saved.
- The structures resulting from the ABCluster calculations were further optimized with GFN1-xTB using the xtb program.
- After filtering out fragmented/reacted structures and structures 50 kcal mol^{-1} higher in electronic energy than the lowest minima found, 5 structures were randomly selected for each cluster.

This resulted in a set of 1929 structures (slightly less than $5 \times 395 = 1975$ due to some cluster having less than 5 unique equilibrium structures). We subsequently removed an additional ~100 structures that later exhibited geometry optimization convergence issues, resulting in a final dataset consisting of 1,831 non-equilibrium structures.

All structures were optimized at the ωB97X-D/6-31++G(d,p) level of theory. If the optimization of a structure failed (e.g., convergence failure or remaining imaginary frequencies), it was retried up to two times. Structures that still did not properly optimize were removed from the dataset. In the end, 1,831 structures remained, which were then used as the starting point for geometry optimizations with the quantum chemistry methods included in the benchmark.

Microhydrated (sulfuric acid–ammonia)-pair clusters:

For each of the $(SA_1AM_1)_{1-6}W_{0-10}$ clusters, we applied the following procedure:

- In three independent simulations, 300 guess structures were generated with the rigid-molecule optimizer in ABCluster.[2,3] This structure set was optimized over 100 steps, and the lowest 300 structures were saved.
- These were further optimized with GFN1-xTB using the xtb program.
- After filtering out fragmented/reacted structures, 3 structures were uniformly selected for each cluster from the lowest 50 kcal mol^{-1} minima.

Hydrated sulfuric acid–ammonia ($SA_{0-3}AM_{0-3}W_{0-5}$),
sulfuric acid–dimethylamine ($SA_{0-3}DMA_{0-3}W_{0-5}$), and
methanesulfonic acid–methylamine ($MSA_{0-3}MA_{0-3}W_{0-5}$) *clusters*

To obtain the lowest free energy conformers we followed the following procedure:

- In five independent simulations, 200×(*number-of-monomers*) guess structures were generated with the rigid-molecule optimizer in ABCluster.[2,3] These structure were optimized over 100 steps, and the lowest 2,000 structures were saved.
- These were further optimized with GFN1-xTB using the xtb program.
- After filtering out fragmented/reacted structures, 1,000 were uniformly selected for each cluster based on radius of gyration, electronic energy, and dipole moment.
- All these structures were further optimized at the B97-3c level of theory.
- 100 unique and non-exploded structures within 10 kcal mol^{-1} of the lowest electronic energy minimum were taken to the next step.
- These structures were optimized at ωB97X-D/6-31++G(d,p) followed by a vibrational analysis. If the optimization did not converge or we obtained imaginary vibrational frequencies, we repeated this step up to two times. Structures that still did not converge without imaginary frequencies were removed from the dataset.
- The lowest structure was refined at DLPNO$^{\text{NormalPNO}}$–CCSD(T_0)/aug-cc-pVTZ.

S2 Supporting Data and Structures

The coordinates of the calculated structures and the associated thermochemistry are available in the Atmospheric Cluster Database at:

https://github.com/elmjonas/ACDB.git

The datasets for this work can be found under the Articles/neefjes25_hydration folder. The folder contains all data in corresponding subfolder:

- el_energy_benchmark
- el_energy_error_scaling
- opt_geometries
- vibrational_frequencies
- hydration_distribution
- binding_free_energies

We use the JK framework to pickle our databases. It is a free, open-source program available at:

https://github.com/kubeckaj/JKCS2.1

The package contains Bash and Python codes and is suitable for any GNU/Linux system. Each script contains its own help function, and the overall manual with recommended approaches is available at:

https://jkcs.readthedocs.io

S3 Particle Formation Rate Calculations

We used the Atmospheric Cluster Dynamics Code (ACDC) to calculate particle formation rates J. ACDC constructs and solves the system's birth-death equations, which describe the cluster population dynamics.[4] Collision rate coefficients (β) were obtained from kinetic gas theory,[4] while evaporation rate coefficients (γ) were derived from quantum chemistry (QC) calculations via detailed balance:

$$\gamma[(x+y) \rightarrow x+y] = \beta[x+y \rightarrow (x+y)] \cdot \frac{p_{\mathrm{ref}}}{RT} \exp\left(\frac{\Delta G^{\ominus}[x+y] - \Delta G^{\ominus}[x] - \Delta G^{\ominus}[y]}{RT}\right), \quad (1)$$

where R is the gas constant, T the temperature, p_{ref} = 1 atm a reference pressure, and $\Delta G^{\ominus}$ the binding free energy.

We evaluated J over the relative humidity (RH) range 0–100% at 278.15 K and 298.15 K for three systems:

- $\mathrm{MSA}_{0-3}\mathrm{MA}_{0-3}\mathrm{W}_{0-5}$
- $\mathrm{SA}_{0-3}\mathrm{AM}_{0-3}\mathrm{W}_{0-5}$
- $\mathrm{SA}_{0-3}\mathrm{DMA}_{0-3}\mathrm{W}_{0-5}$

All monomer concentrations were kept constant, assuming that cluster formation does not significantly deplete monomers. The following monomer concentrations were used: SA and MSA at 10^6, 10^7, and 10^8 cm^{-3}; AM at 10 and 10,000 ppt; DMA at 1 and 10 ppt; and MA at 1 and 10 ppt, covering typical boundary-layer ranges.

Binding free energies were computed at the DLPNO$^{\mathrm{TightPNO}}$-CCSD(T$_0$)/aug-cc-pVTZ//ωB97X-D/6-31++G(d,p) level of theory with quasi-harmonic corrections applied. In addition, we also corrected the binding free energies with $-N_{\mathrm{vib}}k_{\mathrm{B}}T/4$ according to the Halonen limit. Thus, we considered two limiting cases: (1) the QC calculations without any harmonic corrections, and (2) the largest corrections that can be expected.

The calculated J dependents on the allowed outgrowing pathways. In ACDC, outgrowing paths are defined by cluster compositions outside the simulated clusters that

undergo no evaporation back into the simulated system. The outgrowing paths should include the cluster compositions for which the cluster growth is most likely to spontaneously continue towards a particle. Previous studies have shown that sulfuric acid–base clusters are most stable when there is an approximately 1:1 acid–base ratio.[5,6] For the systems studied here, we, therefore, used outgrowing paths $(\text{acid})_{3+1}(\text{base})_3$, $(\text{acid})_3(\text{base})_{3+1}$, and $(\text{acid})_{3+1}(\text{base})_{3+1}$. If the system grows outside of the system in any other direction, the resulting cluster is considered unstable and is immediately evaporated back to the nearest simulation cluster.

Because the maximum simulated cluster size was relatively small, the critical cluster size, where growth starts to vastly outweigh evaporation, might not be included. When the critical cluster is not in the simulated system, the obtained particle formation rate will be larger than the actual particle formation rate J, as outgrowing paths are defined as having no evaporation back into the system. Elm et al.[6] coined the term: potential particle formation rate $J_{\text{potential}}$, to indicate that the results cannot directly be related to the actual particle formation rate J, but are rather a measure for the importance of different compounds in cluster formation.

Coagulation loss (CL) of the clusters was included using an exponential loss function

$$\text{CL} = \text{CL}_{\text{ref}} \left(\frac{d}{d_{\text{ref}}} \right)^m \tag{2}$$

where CL_{ref} is a chosen reference coagulation sink, d the cluster diameter, d_{ref} a reference diameter, and m a parameter dependent on the scavenger distribution. We have used $\text{CS}_{\text{ref}} = 10^{-3}$, d_{ref} equal to the diameter of SA, and $m = -1.6$, which corresponds to typical coagulation loss in the boreal forest environment.[7,8]

In total, we examined J as a function of RH for three different systems, with and without Halonen-limit corrections, at two temperatures, three acid concentrations, and two base concentrations, amounting to 72 distinct ACDC simulations.

S4 Electronic Binding Energy Benchmark Results

Figure S1 shows the signed errors of electronic binding energies for the benchmarked QC methods compared to two reference methods.

Figure S2 shows the signed errors per electron of electronic binding energies for the benchmarked QC methods compared to two reference methods.

Table S2 provides mean absolute errors and mean errors of electronic binding energies per cluster and per electron for the microhydrated monomers and dimers, using the benchmarked QC methods with CCSD(T_0)-F12/cc-pVTZ-F12 as reference method.

Table S1 provides mean absolute errors and mean errors of electronic binding energies per cluster and per electron for the microhydrated monomers and dimers, using the benchmarked QC methods with DLPNO$^{\mathrm{NormalPNO}}$-CCSD(T_0)/aug-cc-pVTZ as reference method.

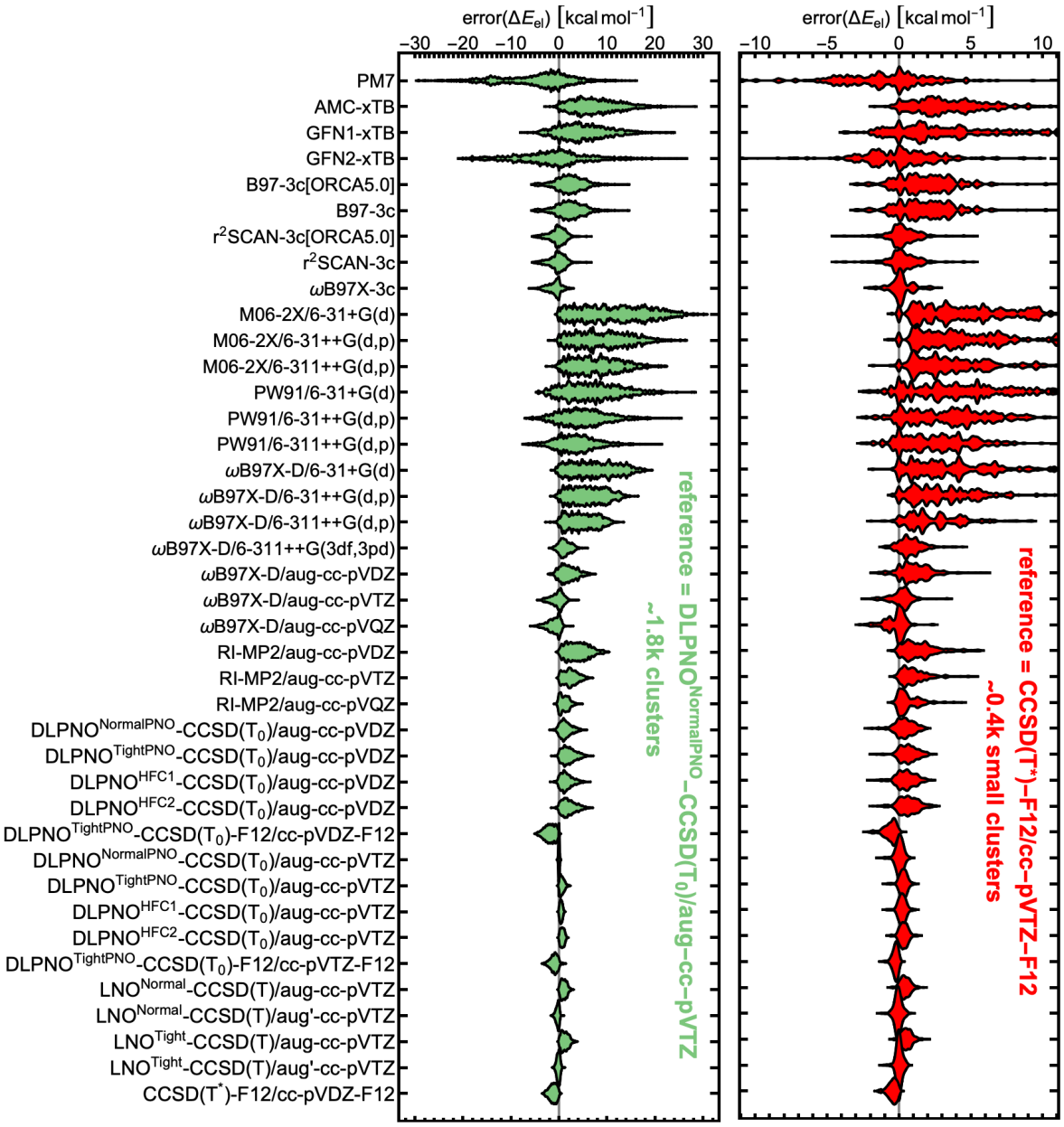

Figure S1: Signed errors of electronic binding energies for up to five configurations of the studied microhydrated monomers and dimers, using the benchmarked quantum chemistry (QC) methods and compared to two different reference methods. The microhydrated monomers and dimers consist of all 395 unique combinations of molecules that follow the chemical formula $(\text{acid and/or base})_{0\text{–}2}\text{W}_{0\text{–}5}$.

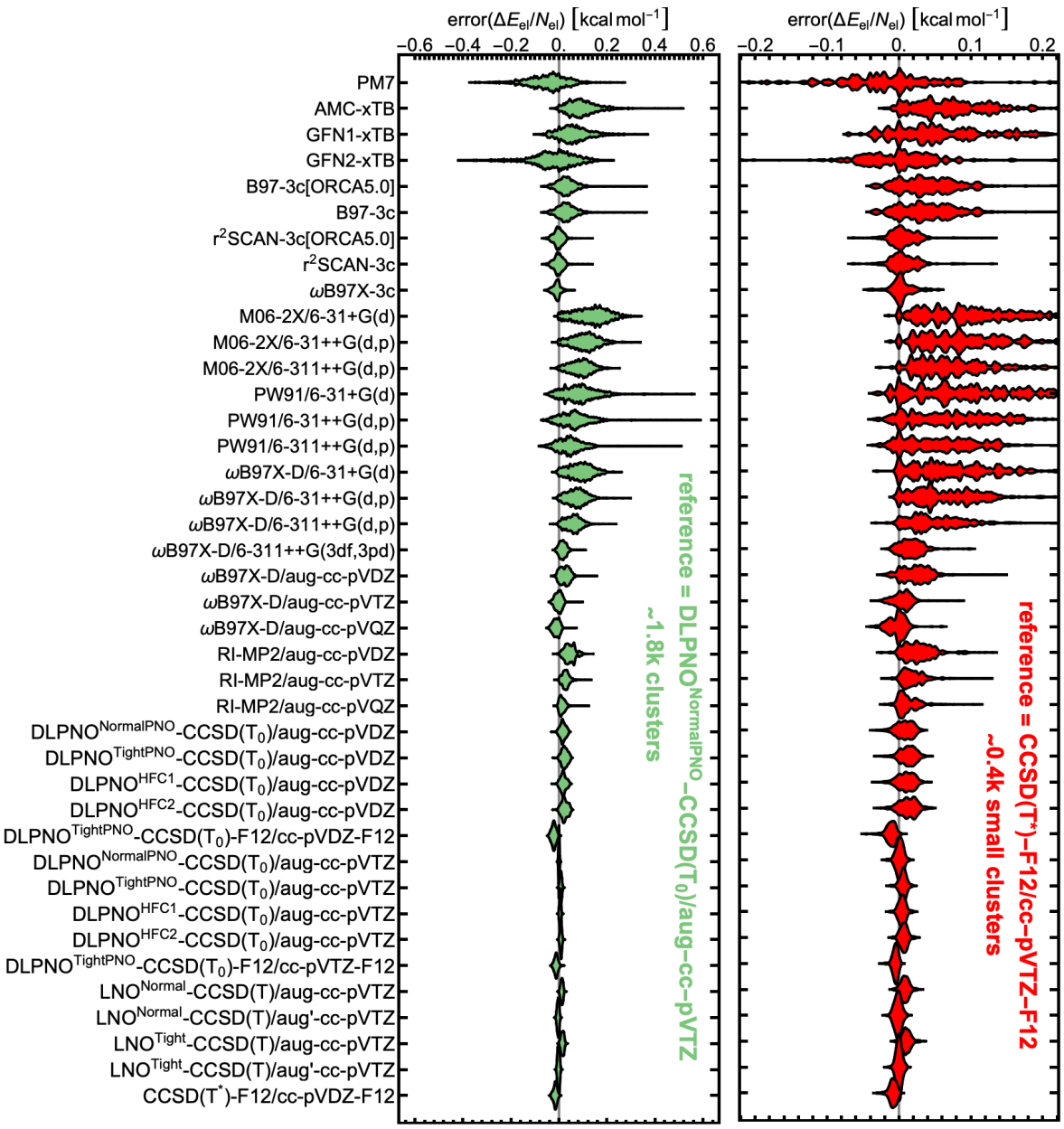

Figure S2: Signed errors per electron of electronic binding energies for up to five configurations of the studied microhydrated monomers and dimers, using the benchmarked quantum chemistry (QC) methods and compared to two different reference methods. The microhydrated monomers and dimers consist of all 395 unique combinations of molecules that follow the chemical formula (acid and/or base)$_{0–2}$W$_{0–5}$.

Table S1: El. energy errors compared to CCSD(T*)-F12/cc-pVTZ-F12.

method	$\langle\lvert E_{el}\rvert\rangle$	$\langle E_{el}\rangle$	$\langle\lvert E_{el}/N_{el}\rvert\rangle$	$\langle E_{el}/N_{el}\rangle$
PM7	+3.12	−1.64	+0.05809	−0.02233
AMC-xTB	+4.04	+4.01	+0.00520	+0.00520
GFN1-xTB	+3.89	+3.50	+0.08290	+0.08243
GFN2-xTB	+2.38	−0.40	+0.00062	+0.00062
B97-3c[ORCA5.0]	+2.24	+1.96	+0.07620	+0.06809
B97-3c	+2.24	+1.95	+0.00292	+0.00292
r^2SCAN-3c[ORCA5.0]	+0.75	+0.28	+0.04703	−0.00995
r^2SCAN-3c	+0.75	+0.28	+0.00089	+0.00089
ωB97X-3c	+0.53	+0.16	+0.04638	+0.04202
M06-2X/6-31+G(d)	+5.49	+5.49	+0.14749	+0.14749
M06-2X/6-31++G(d,p)	+4.59	+4.59	+0.04634	+0.04184
M06-2X/6-311++G(d,p)	+3.72	+3.71	+0.14544	+0.14544
PW91/6-31+G(d)	+4.64	+4.50	+0.01464	+0.00644
PW91/6-31++G(d,p)	+3.93	+3.79	+0.18273	+0.18273
PW91/6-311++G(d,p)	+3.06	+2.88	+0.01464	+0.00645
ωB97X-D/6-31+G(d)	+3.66	+3.65	+0.16102	+0.16102
ωB97X-D/6-31++G(d,p)	+3.16	+3.15	+0.01028	+0.00364
ωB97X-D/6-311++G(d,p)	+2.46	+2.42	+0.57391	+0.57391
ωB97X-D/6-311++G(3df,3pd)	+0.90	+0.79	+0.10916	+0.10908
ωB97X-D/aug-cc-pVDZ	+1.24	+1.12	+0.51834	+0.51834
ωB97X-D/aug-cc-pVTZ	+0.55	+0.17	+0.09137	+0.09125
ωB97X-D/aug-cc-pVQZ	+0.51	−0.22	+0.72390	+0.72390
RI-MP2/aug-cc-pVDZ	+1.49	+1.47	+0.07387	+0.07363
RI-MP2/aug-cc-pVTZ	+1.07	+1.06	+0.96982	+0.96982
RI-MP2/aug-cc-pVQZ	+0.69	+0.66	+0.09477	+0.09272
DLPNO$^{\text{NormalPNO}}$-CCSD(T_0)/aug-cc-pVDZ	+0.61	+0.40	+0.32252	+0.32252
DLPNO$^{\text{TightPNO}}$-CCSD(T_0)/aug-cc-pVDZ	+0.74	+0.65	+0.08094	+0.07875
DLPNO$^{\text{HFC1}}$-CCSD(T_0)/aug-cc-pVDZ	+0.66	+0.50	+0.46430	+0.46430
DLPNO$^{\text{HFC2}}$-CCSD(T_0)/aug-cc-pVDZ	+0.77	+0.66	+0.06355	+0.06079
DLPNO$^{\text{TightPNO}}$-CCSD(T_0)-F12/cc-pVDZ-F12	+0.61	−0.61	+0.59209	+0.59209
DLPNO$^{\text{NormalPNO}}$-CCSD(T_0)/aug-cc-pVTZ	+0.20	+0.05	+0.07470	+0.07441
DLPNO$^{\text{TightPNO}}$-CCSD(T_0)/aug-cc-pVTZ	+0.34	+0.30	+0.38612	+0.38612
DLPNO$^{\text{HFC1}}$-CCSD(T_0)/aug-cc-pVTZ	+0.28	+0.22	+0.06489	+0.06474
DLPNO$^{\text{HFC2}}$-CCSD(T_0)/aug-cc-pVTZ	+0.37	+0.34	+0.57357	+0.57357
DLPNO$^{\text{TightPNO}}$-CCSD(T_0)-F12/cc-pVTZ-F12	+0.25	−0.25	+0.05082	+0.05030
LNO$^{\text{Normal}}$-CCSD(T)/aug-cc-pVTZ	+0.46	+0.44	+0.81668	+0.81668
LNO$^{\text{Normal}}$-CCSD(T)/aug'-cc-pVTZ	+0.19	−0.06	+0.01851	+0.01691
LNO$^{\text{Tight}}$-CCSD(T)/aug-cc-pVTZ	+0.56	+0.55	+5.33980	+5.33980
LNO$^{\text{Tight}}$-CCSD(T)/aug'-cc-pVTZ	+0.15	+0.02	+0.02545	+0.02366
CCSD(T*)-F12/cc-pVDZ-F12	+0.45	−0.45	+1.70643	+1.70643

Table S2: El. energy errors compared to DLPNO$^{\text{NormalPNO}}$-CCSD(T_0)/aug-cc-pVTZ.

method	$\langle\lvert E_{el}\rvert\rangle$	$\langle E_{el}\rangle$	$\langle\lvert E_{el}/N_{el}\rvert\rangle$	$\langle E_{el}/N_{el}\rangle$
PM7	+6.39	−4.70	+0.07233	−0.04866
AMC-xTB	+7.79	+7.77	+0.00520	+0.00520
GFN1-xTB	+5.77	+5.19	+0.09551	+0.09527
GFN2-xTB	+5.73	−1.22	+0.00062	+0.00062
B97-3c[ORCA5.0]	+3.00	+2.33	+0.07050	+0.06227
B97-3c	+3.00	+2.33	+0.00292	+0.00292
r^{2}SCAN-3c[ORCA5.0]	+1.29	−0.32	+0.06709	−0.02110
r^{2}SCAN-3c	+1.29	−0.32	+0.00089	+0.00089
ωB97X-3c	+1.26	−0.94	+0.03850	+0.03108
M06-2X/6-31+G(d)	+11.56	+11.56	+0.14749	+0.14749
M06-2X/6-31++G(d,p)	+9.18	+9.18	+0.03852	+0.03103
M06-2X/6-311++G(d,p)	+7.76	+7.74	+0.14544	+0.14544
PW91/6-31+G(d)	+7.36	+7.06	+0.01546	−0.00233
PW91/6-31++G(d,p)	+5.71	+5.15	+0.18273	+0.18273
PW91/6-311++G(d,p)	+4.33	+3.54	+0.01546	−0.00233
ωB97X-D/6-31+G(d)	+7.51	+7.50	+0.16102	+0.16102
ωB97X-D/6-31++G(d,p)	+6.07	+6.05	+0.01445	−0.00932
ωB97X-D/6-311++G(d,p)	+4.90	+4.88	+0.57391	+0.57391
ωB97X-D/6-311++G(3df,3pd)	+1.37	+1.21	+0.13387	+0.13380
ωB97X-D/aug-cc-pVDZ	+2.06	+1.91	+0.51834	+0.51834
ωB97X-D/aug-cc-pVTZ	+0.99	−0.31	+0.10699	+0.10687
ωB97X-D/aug-cc-pVQZ	+1.25	−1.08	+0.72390	+0.72390
RI-MP2/aug-cc-pVDZ	+3.87	+3.87	+0.08932	+0.08910
RI-MP2/aug-cc-pVTZ	+2.41	+2.41	+0.96982	+0.96982
RI-MP2/aug-cc-pVQZ	+1.17	+1.14	+0.09084	+0.08691
DLPNO$^{\text{NormalPNO}}$-CCSD(T_0)/aug-cc-pVDZ	+1.48	+1.41	+0.32252	+0.32252
DLPNO$^{\text{TightPNO}}$-CCSD(T_0)/aug-cc-pVDZ	+2.09	+2.07	+0.07194	+0.06543
DLPNO$^{\text{HFC1}}$-CCSD(T_0)/aug-cc-pVDZ	+1.62	+1.58	+0.46430	+0.46430
DLPNO$^{\text{HFC2}}$-CCSD(T_0)/aug-cc-pVDZ	+2.07	+2.05	+0.05487	+0.04599
DLPNO$^{\text{TightPNO}}$-CCSD(T_0)-F12/cc-pVDZ-F12	+1.91	−1.91	+0.59209	+0.59209
DLPNO$^{\text{NormalPNO}}$-CCSD(T_0)/aug-cc-pVTZ	+0.00	+0.00	+0.08907	+0.08882
DLPNO$^{\text{TightPNO}}$-CCSD(T_0)/aug-cc-pVTZ	+0.66	+0.65	+0.38612	+0.38612
DLPNO$^{\text{HFC1}}$-CCSD(T_0)/aug-cc-pVTZ	+0.44	+0.44	+0.07293	+0.07273
DLPNO$^{\text{HFC2}}$-CCSD(T_0)/aug-cc-pVTZ	+0.70	+0.70	+0.57357	+0.57357
DLPNO$^{\text{TightPNO}}$-CCSD(T_0)-F12/cc-pVTZ-F12	+1.06	−1.02	+0.05845	+0.05818
LNO$^{\text{Normal}}$-CCSD(T)/aug-cc-pVTZ	+1.03	+1.03	+0.81668	+0.81668
LNO$^{\text{Normal}}$-CCSD(T)/aug'-cc-pVTZ	+0.34	−0.30	+0.01719	+0.01541
LNO$^{\text{Tight}}$-CCSD(T)/aug-cc-pVTZ	+1.26	+1.26	+5.33980	+5.33980
LNO$^{\text{Tight}}$-CCSD(T)/aug'-cc-pVTZ	+0.28	−0.13	+0.02611	+0.02446
CCSD(T*)-F12/cc-pVDZ-F12	+1.26	−1.26	+1.70643	+1.70643

S5 Equilibrium Cluster Geometry Comparisons

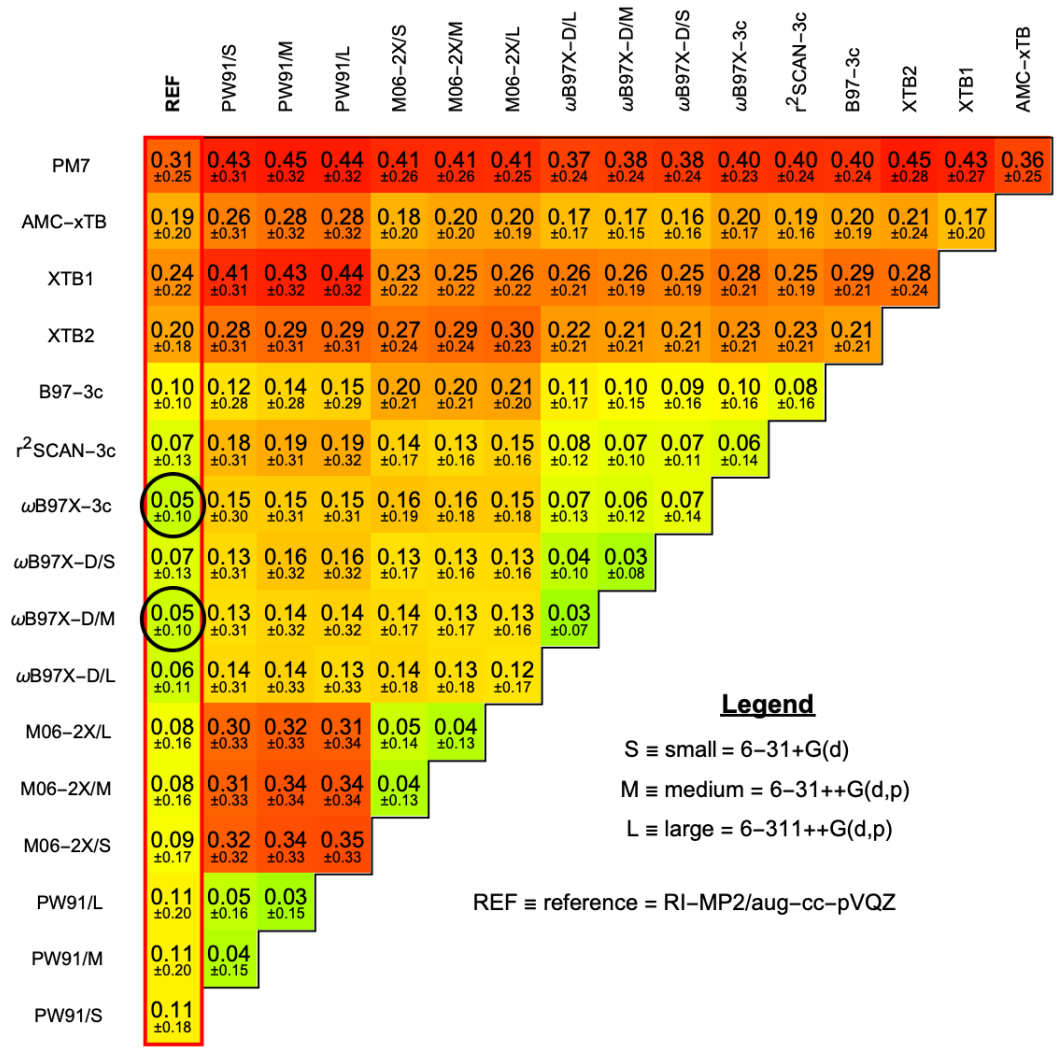

Figure S3: Mean of root-mean-square deviations (RMSD [Å]) between (acid and/or base)$_{0–2}$W$_{0–5}$ equilibrium geometries optimized at different levels of theory. Black circles indicate the best-performing methods relative to the reference, RI-MP2/aug-cc-pVQZ (REF).

S6 Vibrational Frequency Analysis

To investigate the vibrational contributions to the thermal motion of hydrated molecular systems, we calculated the vibrational frequencies for all (acid and/or base)$_{0-2}$W$_{0-5}$ clusters using 5 different QC methods: RI-MP2/aug-cc-pVQZ (~0.6k clusters), ωB97X-D/6-31++G(d,p) (~1.8k), ωB97X-3c (~1.8k), r^{2}SCAN-3c (~1.8k), and B97-3c (~1.8k). Vibrational analysis is computationally demanding, therefore these methods were selected for their efficiency and their relatively strong performance in Secs. 3.1 and 3.2.

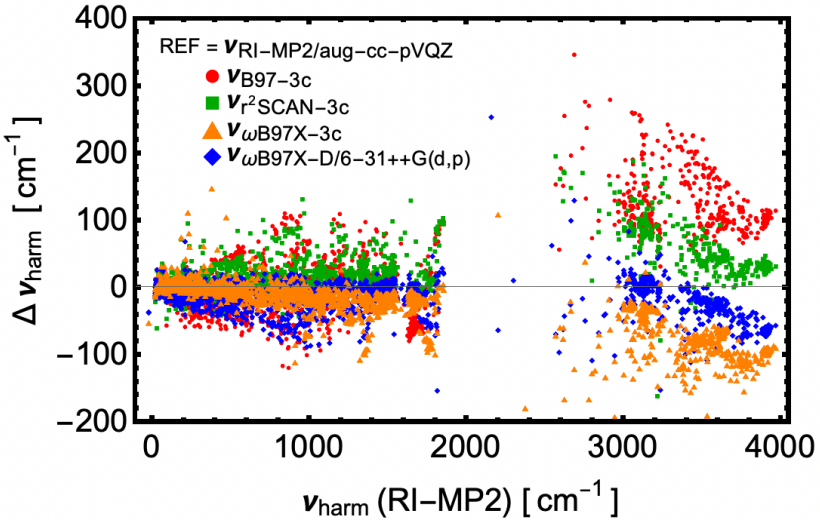

Figure S4: The differences in harmonic vibrational frequencies (cm^{-1}) calculated using the B97-3c, r^{2}SCAN-3c, ωB97X-3c, and ωB97X-D/6-31++G(d,p) methods compared to those from the RI-MP2/aug-cc-pVQZ method, plotted against the harmonic vibrational frequencies (cm^{-1}) at the RI-MP2/aug-cc-pVQZ method. The histogram above illustrates the distribution of vibrational frequencies for monomers (red) and clusters (black) at the ωB97X-3c level of theory.

Initially, we used the (rigid-rotor) harmonic oscillator approximation without any correction. We are not aware of any computational method that consistently provides accurate vibrational frequencies at feasible computational costs for molecular clusters. As RI-MP2 has been shown to provide good equilibrium structures in previous studies,[9,10] we used it as a reference. It is important to note that we do not claim that RI-MP2/aug-cc-pVQZ provides correct vibrational frequencies. Rather, we are interested

in comparing the vibrational frequencies between methods to see the spread in values. Figure S4 shows how the harmonic vibrational frequencies of ωB97X-D/6-31++G(d,p), ωB97X-3c, r^{2}SCAN-3c, and B97-3c differ compared to those of the RI-MP2/aug-cc-pVQZ method. Vibrations under 2,000 cm^{-1} seem to be predicted similarly across the five methods. Above 2,000 cm^{-1}, there are, however, significant differences between the methods. For each frequency in this region, there is a discrepancy of several 100 wavenumbers, with each method having a near systematic deviation across vibrational frequencies compared to the others. As the number of vibrational frequencies increases with cluster size, these discrepancies will result in increasingly larger errors for the vibrational contributions to the system's thermodynamics. Proton transfers within clusters lead to the emergence of new peaks and shifts in the spectra compared to monomers. This, combined with the abundance of low vibrational frequencies in clusters, has the greatest impact on the accuracy of the binding free energy.

S6.1 Low-vibrational frequencies

Figure S5 shows the thermal energy and entropic contributions of the vibrational frequencies (kcal mol^{-1}) as a function of their wavenumber (cm^{-1}) according to the harmonic oscillator approximation. Above 800 cm^{-1}, the entropic contribution is essentially zero, while the thermal energy contribution increases linearly with the wavenumber. Hence, errors in vibrational frequencies belonging to this region will have a similar effect on the free energy regardless of the magnitude of the frequency. As the vibrational frequencies tend towards zero, the entropic contribution tends towards $-\infty$. This is an artifact of the harmonic oscillator treatment. Low vibrational frequencies are often present in molecular clusters. In the quasi-harmonic approximation,[11] their vibrational entropy contributions are replaced by rotational entropy. When calculating the vibrational entropy values of the clusters studied here, we used a rotor-vibration crossover

frequency of 100 cm^{-1} under which the quasi-harmonic approximation is used. Note that some programs (e.g., ORCA) use the quasi-harmonic approximation by default.

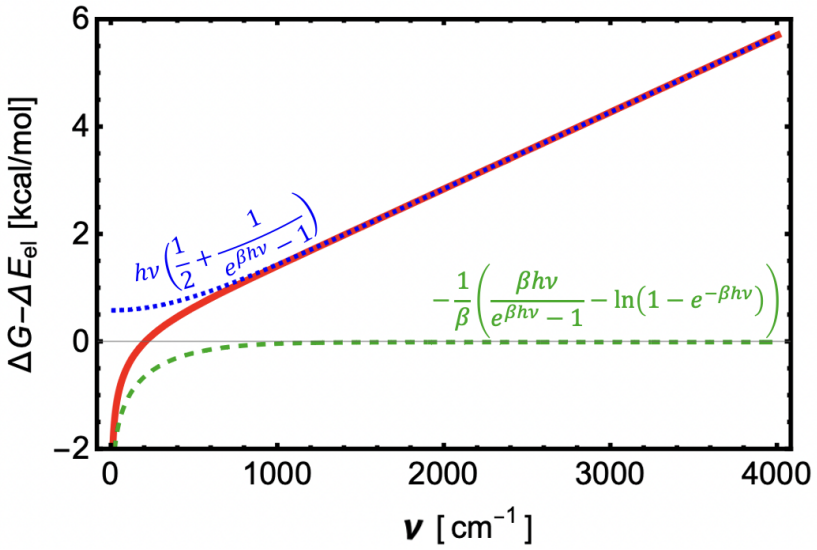

Figure S5: The thermal enthalpy correction (blue dotted line), entropic contribution ($-TS$; green dashed line), and total thermal Gibbs free energy contribution (red line) of vibrations as functions of the wavenumber (cm^{-1}).

S6.2 Anharmonic frequencies

Anharmonicity can be incorporated using methods such as second-order vibrational perturbation theory (VPT2[12]) with a significant increase in computational cost. As a more practical alternative, vibrational scaling factors are often applied to correct for anharmonic effects. Jacobsen et al.[13] further note that anharmonic vibrational frequency calculations are not worthwhile when small basis sets are used, as the error introduced by the basis set exceeds that from neglecting anharmonicity, making scaling factors sufficient. The Computational Chemistry Comparison and Benchmark Database (CCCBDB) provides scaling factors for various QC methods to address anharmonicity.[14] For instance, the scaling factor for ωB97X-D/6-31++G(d,p) is listed as 0.952, while that for MP2/aug-cc-pVQZ is 0.950. In contrast to the CCCBDB, Myllys et al.[15] derived for a small set of molecular clusters a scaling factor of 0.996 (nearly unity) for

the ωB97X-D/6-31++G(d,p) method. However, at the time of this study, no scaling factors were available for DFT-3c methods. Therefore, in this work we calculate our own scaling factors.

We performed VPT2 calculations for ωB97X-D/6-31++G(d,p), ωB97X-3c, r^{2}SCAN-3c, and B97-3c, excluding RI-MP2/aug-cc-pVQZ due to its high computational cost. These calculations are computationally intensive, requiring accurate equilibrium geometries (achieved by *VeryTightOpt* criteria) and strict SCF settings (achieved by *ExtremeSCF* criteria). Using less precise settings introduces significant numerical errors, sometimes leading to unphysical data. To address this, we removed all frequencies where the anharmonic frequency becomes imaginary.

Figures S6, S7, S8, and S9 show the correlation of the anharmonic vibrational frequencies (ν_{anh}) and the corresponding harmonic frequencies (ν_{harm}) as well as the ratio between the two. All frequencies are shown as gray data points. To reduce noise, the median over 20 data points is represented by black points, while the red line shows the median over 100 cm^{-1} bins. From the left side, we observe that most vibrational frequencies are located below 2000 cm^{-1} and above 3000 cm^{-1}, with median harmonic frequencies closely following anharmonic ones. However, the spread of the ration is quite significant for small vibrational frequencies. Also, in the range between 2000 and 3000 cm^{-1}, fewer vibrational frequencies are present, and the discrepancy between anharmonic and harmonic frequencies is notably larger. These frequencies mainly correspond to O–H and possibly N–H vibrations, which exhibit a higher degree of anharmonicity. Given the complex variation of the $\nu_{\mathrm{anh}}/\nu_{\mathrm{harm}}$ ratio with ν_{harm}, we decided to explore multiple scaling corrections for the anharmonicity.

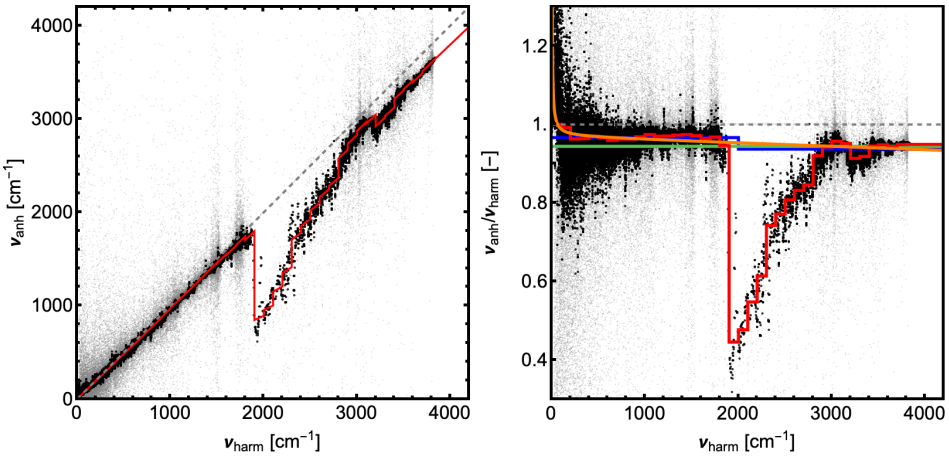

Figure S6: Left: The anharmonic vibrational frequency (ν_{anh}; cm^{-1}) as a function of the corresponding harmonic vibrational frequency (ν_{harm}; cm^{-1}) at the B97-3c. Gray points represent all vibrational frequencies. The black points show medians of the 20 nearest frequencies. Red line show binned medians over 100 cm^{-1} intervals. Right: The ratio $\nu_{\text{anh}}/\nu_{\text{harm}}$ as a function of ν_{harm} (cm^{-1}) with single scaling factor (green), two-region scaling (blue), multi-region scaling (red), and scaling function (orange) shown.

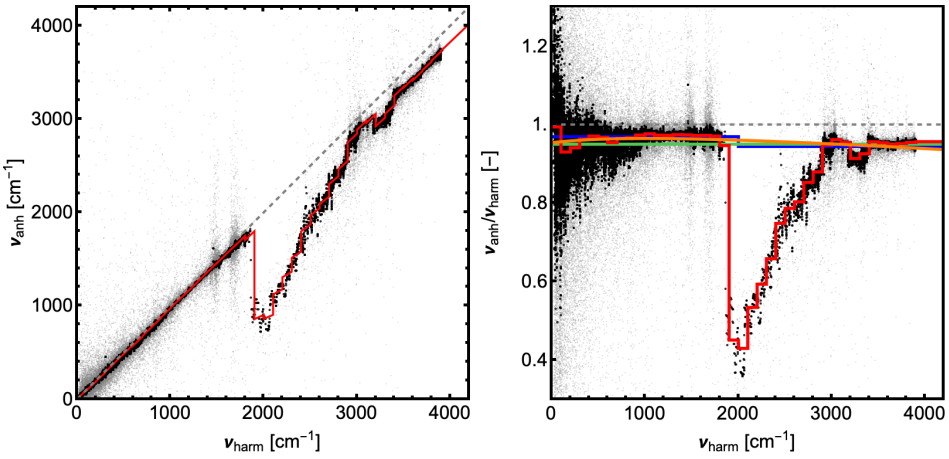

Figure S7: Left: The anharmonic vibrational frequency (ν_{anh}; cm^{-1}) as a function of the corresponding harmonic vibrational frequency (ν_{harm}; cm^{-1}) at the r^{2}SCAN-3c. Gray points represent all vibrational frequencies. The black points show medians of the 20 nearest frequencies. Red line show binned medians over 100 cm^{-1} intervals. Right: The ratio $\nu_{\text{anh}}/\nu_{\text{harm}}$ as a function of ν_{harm} (cm^{-1}) with single scaling factor (green), two-region scaling (blue), multi-region scaling (red), and scaling function (orange) shown.

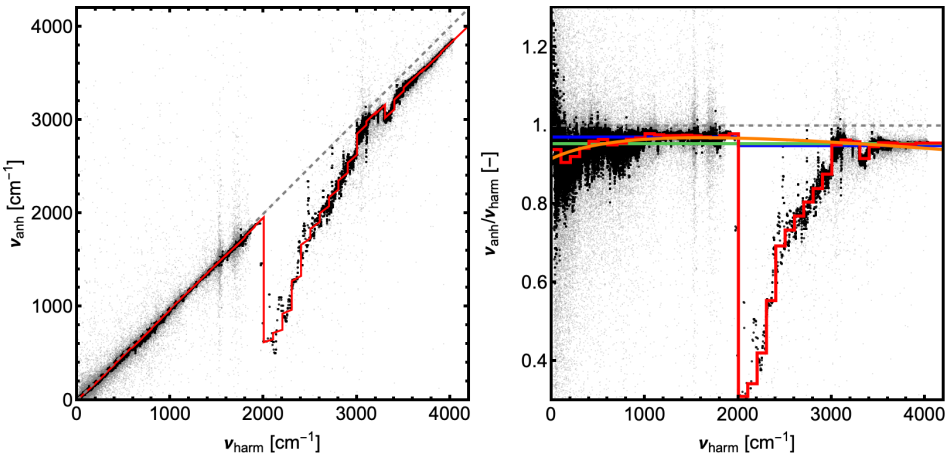

Figure S8: Left: The anharmonic vibrational frequency (ν_{anh}; cm^{-1}) as a function of the corresponding harmonic vibrational frequency (ν_{harm}; cm^{-1}) for the ωB97X-3c, with VPT2 calculations performed using ExtremeSCF settings. Gray points represent all vibrational frequencies. The black points and red line show medians of the five nearest vibrational frequencies and binned medians over 100 cm^{-1} intervals, respectively, with frequencies where $\nu_{\text{anh}} > \nu_{\text{harm}}$ removed. Right: The ratio $\nu_{\text{anh}}/\nu_{\text{harm}}$ as a function of ν_{harm} (cm^{-1}). The single scaling factor (green), two-region scaling factor (blue), and scaling function (orange) are also shown.

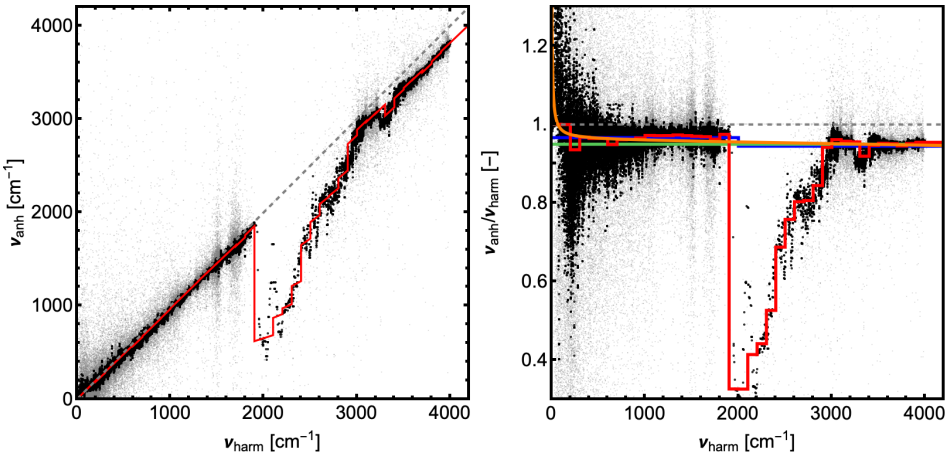

Figure S9: Left: The anharmonic vibrational frequency (ν_{anh}; cm^{-1}) as a function of the corresponding harmonic vibrational frequency (ν_{harm}; cm^{-1}) for the ωB97X-D/6-31++G(d,p), with VPT2 calculations performed using ExtremeSCF settings. Gray points represent all vibrational frequencies. The black points and red line show medians of the five nearest vibrational frequencies and binned medians over 100 cm^{-1} intervals, respectively, with frequencies where $\nu_{\text{anh}} > \nu_{\text{harm}}$ removed. Right: The ratio $\nu_{\text{anh}}/\nu_{\text{harm}}$ as a function of ν_{harm} (cm^{-1}). The single scaling factor (green), two-region scaling factor (blue), and scaling function (orange) are also shown.

S6.3 Harmonic frequency scaling

We explored four different scaling approaches of harmonics vibrational frequencies:

1. A single constant scaling factor across the entire frequency range
2. Two separate constant scaling factors for the regions below and above 2000 cm^{-1}
3. Multi-region constant scaling factors with a region size of 100 cm^{-1}
4. An arbitrary scaling function of the form $A - B \cdot \nu_{\text{harm}} - C/(D + \nu_{\text{harm}})$

The coefficients for these scaling approaches, excluding the multi-region approach, for the different QC methods are presented in Table S3. For the single constant scaling factors, values range from 0.944 to 0.954. The scaling factor of 0.950 for ωB97X-D/6-31++G(d,p) is in reasonable agreement with the value of 0.952 from the CCCBDB, but differs from the 0.996 scaling factor proposed by Myllys et al.[15] that was obtained only for a (likely too) small set of molecular clusters. All scaling factors/function are also visualized for each method in Figs S6, S7, S8, and S9.

Table S3: The single scaling factor (f), two-region scaling factors ($f_<$ and $f_>$) for the frequency ranges below and above 2000 cm^{-1}, and the coefficients (A, B, C, D) for the scaling function, obtained by fitting to the difference between the anharmonic and harmonic vibrational frequencies, calculated using the B97-3c, r^{2}SCAN-3c, ωB97X-3c, and ωB97X-D/6-31++G(d,p) methods.

Methods		B97-3c	r^{2}SCAN-3c	ωB97X-3c	ωB97X-D/ 6-31++G(d,p)
$f \cdot \nu_{\text{harm}}$	f	0.944	0.950	0.954	0.950
$f_< \cdot \nu_{\text{harm}}(< 2000\text{cm}^{-1})$	$f_<$	0.967	0.969	0.971	0.967
$f_\geq \cdot \nu_{\text{harm}}(\geq 2000\text{cm}^{-1})$	$f_\geq$	0.937	0.944	0.949	0.945
$A - B \cdot \nu_{\text{harm}} - \frac{C}{D+\nu_{\text{harm}}}$	A	0.970	1.021	1.076	0.962
	B	0.856×10^{-5}	1.597×10^{-5}	2.428×10^{-5}	0.381×10^{-5}
	C	−1.996	102.5	188.5	−2.814
	D	−0.0448	1517	1179	5.54

S6.4 Experimental frequency benchmark

We compared vibrational frequencies from QC methods with identifiable fundamentals from multiple experimental publications (see Tab. S5). Identifying fundamentals at the low end of experimental spectra is difficult, so the comparisons are biased toward higher wavenumbers. Peak assignment between QC and experiment used the nearest-peak approach. Only observable fundamentals (ν_0 or $\nu_1 \leftarrow \nu_0$) were considered; combination bands and overtones were excluded. Experiments were in the gas phase, with other gases often present, leading to small variations between data.

Table S4: The mean absolute errors (MAE) between experimental (ν^{EXP}) and modeled vibrational frequencies at different QC methods and with different scaling approaches ($S(\nu^{\mathrm{QC}})$) applied. A single scaling factor of 0.950 was used for the RI-MP2 method.[14]

MAE $\nu^{\mathrm{EXP}} - S(\nu^{\mathrm{QC}}_{\mathrm{harm}})$ $[\mathrm{cm}^{-1}]$	B97-3c	r^{2}SCAN-3c	ωB97X-3c	ωB97X-D/ 6-31++G(d,p)	RI-MP2/ aug-cc-pVQZ
—none—	61	86	131	110	97
single scaling factor	81	52	60	39	40
two-region scaling	84	51	58	36	—
multi-region scaling	79	48	181	109	—
scaling function	82	86	54	38	—

MAEs between experimental and modeled vibrational frequencies for the five QC methods, with different scaling approaches, are shown in Tab. S4. Without scaling, B97-3c gives the lowest MAE. However, applying any of the four scaling functions worsens its agreement with experiment. For the other methods, all scaling functions (except multi-region scaling) substantially reduce MAEs compared to the unscaled harmonic approximation. Multi-region scaling performs poorly in the anharmonic region (2000–3000 cm^{-1}). The two-region scaling applied to ωB97X-D/6-31++G(d,p) yields the lowest overall MAE, though the improvement over a single scaling factor is minor. We therefore recommend the single scaling factor as a practical choice.

Table S5: The mean absolute errors (MAE) between experimental vibrational frequencies (ν^{EXP}) and modeled vibrational frequencies at different QC methods with single scaling factor ($S(\nu^{\mathrm{QC}})$) applied. The scaling factors can be found in tab. S3. A single scaling factor of 0.950 was used for the RI-MP2 method.[14] The number in rectangular brackets shows number of obtained experimental frequencies.

MAE $\nu^{\mathrm{EXP}} - S(\nu^{\mathrm{QC}}_{\mathrm{harm}})$ [cm^{-1}]	$\nu^{\mathrm{B97-3c}}_{\mathrm{harm}}$	$S(\nu^{\mathrm{B97-3c}}_{\mathrm{harm}})$	$\nu^{\mathrm{r^2SCAN-3c}}_{\mathrm{harm}}$	$S(\nu^{\mathrm{r^2SCAN-3c}}_{\mathrm{harm}})$	$\nu^{\omega\mathrm{B97-3c}}_{\mathrm{harm}}$	$S(\nu^{\omega\mathrm{B97-3c}}_{\mathrm{harm}})$	$\nu^{\omega\mathrm{B97X-D}}_{\mathrm{harm}}$	$S(\nu^{\omega\mathrm{B97X-D}}_{\mathrm{harm}})$	$\nu^{\mathrm{RI-MP2}}_{\mathrm{harm}}$	$0.95{*}\nu^{\mathrm{RI-MP2}}_{\mathrm{harm}}$
W_1 [2+3+3][16–18]	90	107	143	31	240	91	206	56	162	27
W_2 [3+6+6+12][16,19–21]	90	107	143	31	240	91	206	56	162	27
W_3 [3+9][19,20]	78	118	107	54	206	55	165	32	132	32
W_4 [2][19]	25	174	168	101	172	13	106	13	90	20
W_5 [2+4][19,20]	40	150	60	110	146	52	104	46	95	50
SA_1 [8][22]	73	139	60	109	58	45	40	69	32	63
TMA_1W_1 [3+7][23,24]	129	148	127	127	107	111	121	98	116	114
DMA_1 [1+2][25,26]	54	52	84	15	88	62	119	30	102	17
DMA_1W_1 [3][24]	93	102	105	44	89	103	100	74	85	40
DMA_2 [1][26]	11	64	90	10	138	69	155	20	131	43
MA_1 [12][27]	57	64	72	39	124	44	103	25	108	26
MA_2 [6][27]	70	77	106	25	113	87	112	37	95	29
FA_1 [9+9][28,29]	35	64	48	39	109	39	89	24	67	27
FA_2 [20][28]	23	63	29	44	81	25	72	20	59	24
AM_1 [4][30]	111	50	169	43	171	63	135	28	130	27
AM_2 [3][31]	76	29	143	30	257	91	195	29	183	19
MSA_1 [5][32]	21	99	6	72	76	31	60	34	49	59
MSA_1W_1 [8][32]	51	55	43	41	75	29	56	35	44	28
ACA_1 [16][33]	38	54	47	36	106	29	79	19	70	22
ACA_2 [8][33]	40	20	42	18	57	28	37	12	46	18
NTA_1 [1+9][34,35]	56	57	62	40	132	59	106	37	78	35
NTA_1W_1 [3+5][34,35]	44	86	79	38	168	75	166	55	122	29
NTA_1W_2 [4][34]	74	41	88	58	137	102	113	74	125	89
NTA_2 [8][34]	119	94	126	100	186	105	104	30	131	93
TMA_1 [4+21][25,36]	30	39	30	31	61	24	46	18	50	24
Mean error [cm^{-1}]	61	81	86	52	131	60	110	**39**	97	**40**
Weighted error [cm^{-1}]	58	76	73	50	122	53	102	**38**	88	**38**

S7 Hydration distributions

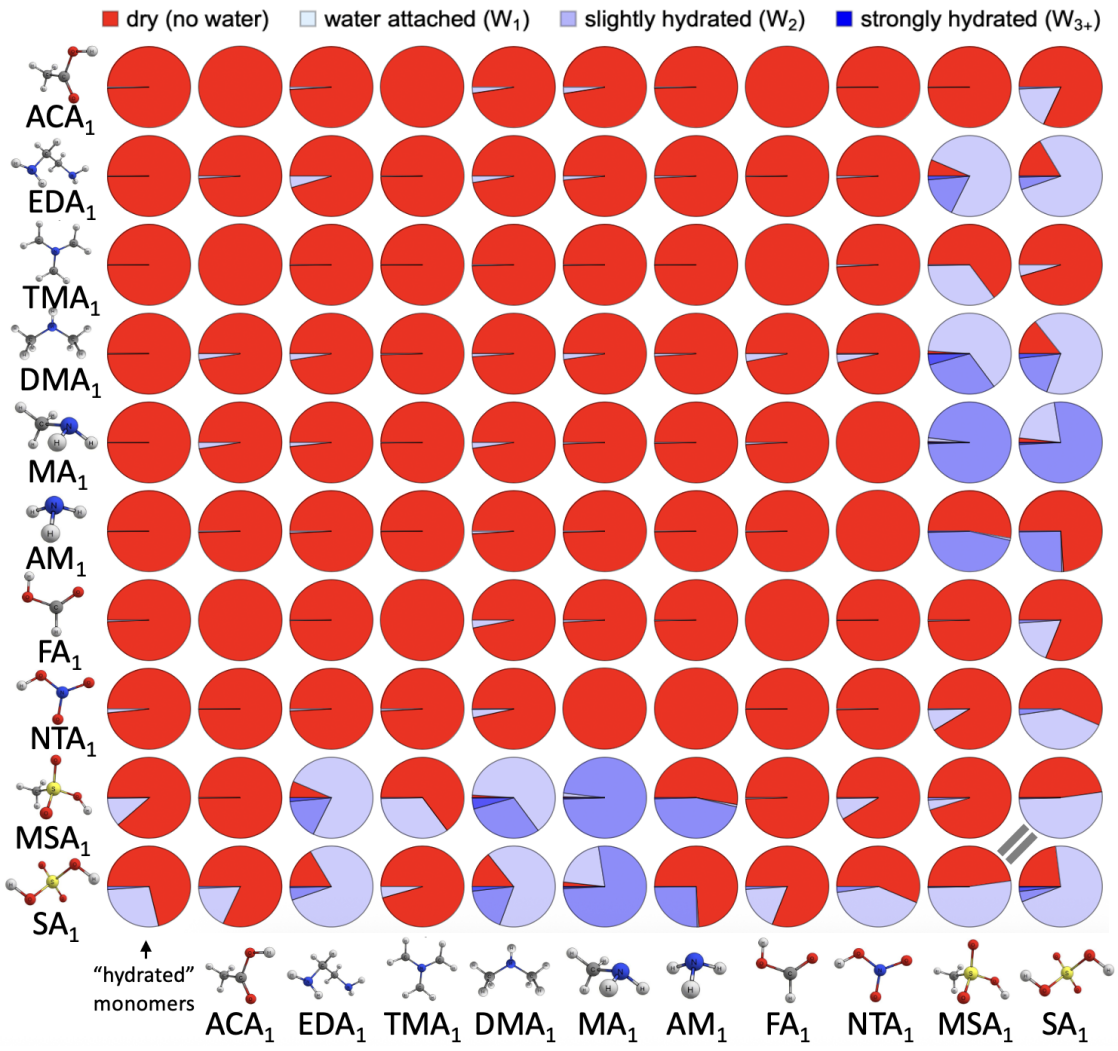

Figure S10: Pie charts of the hydration distribution at 100% relative humidity and 278.15 K for all studied monomers and dimers with up to five water molecules. No hydration is indicated in red, while clusters with one to five water molecules are represented with increasingly darker shades of blue. The results were obtained for the lowest free energy minimum at the DLPNO//DFT level of theory using quasi-harmonic approximation.

S8 New Particle Formation Rates

S8.1 Sulfuric Acid–Dimethylamine

Table S6: Particle formation rates obtained from ACDC simulations.

system	T	SA conc. [cm^{-3}]	base conc. [ppt]	RH	$J^{\mathrm{QC+QHA}}$ [cm^{-3}s^{-1}]	$J^{\mathrm{QC+QHA+HL}}$ [cm^{-3}s^{-1}]
SA-DMA	278.15 K	10^5	1	0%	7.2×10^{-2}	2.5×10^{-2}
SA-DMA	278.15 K	10^5	1	50%	7.6×10^{-2}	4.6×10^{-2}
SA-DMA	278.15 K	10^5	1	100%	1.4×10^{-1}	2.4×10^{-1}
SA-DMA	278.15 K	10^5	10	0%	9.5×10^{-1}	2.7×10^{-1}
SA-DMA	278.15 K	10^5	10	50%	1.	5.1×10^{-1}
SA-DMA	278.15 K	10^5	10	100%	2.2	6.6
SA-DMA	278.15 K	10^7	1	0%	7.4×10^{2}	2.5×10^{2}
SA-DMA	278.15 K	10^7	1	50%	7.9×10^{2}	4.6×10^{2}
SA-DMA	278.15 K	10^7	1	100%	1.6×10^{3}	3.8×10^{3}
SA-DMA	278.15 K	10^7	10	0%	9.3×10^{3}	2.7×10^{3}
SA-DMA	278.15 K	10^7	10	50%	9.9×10^{3}	5.1×10^{3}
SA-DMA	278.15 K	10^7	10	100%	2.1×10^{4}	5.8×10^{4}
SA-DMA	298.15 K	10^5	1	0%	2.2×10^{-6}	1.2×10^{-6}
SA-DMA	298.15 K	10^5	1	50%	2.2×10^{-6}	1.4×10^{-6}
SA-DMA	298.15 K	10^5	1	100%	1.7×10^{-6}	7.6×10^{-7}
SA-DMA	298.15 K	10^5	10	0%	2.1×10^{-3}	1.1×10^{-3}
SA-DMA	298.15 K	10^5	10	50%	2.1×10^{-3}	1.3×10^{-3}
SA-DMA	298.15 K	10^5	10	100%	1.7×10^{-3}	7.1×10^{-4}
SA-DMA	298.15 K	10^7	1	0%	4.	1.8
SA-DMA	298.15 K	10^7	1	50%	3.9	2.
SA-DMA	298.15 K	10^7	1	100%	2.4	$8. \times 10^{-1}$
SA-DMA	298.15 K	10^7	10	0%	4.4×10^{2}	1.2×10^{2}
SA-DMA	298.15 K	10^7	10	50%	4.5×10^{2}	1.8×10^{2}
SA-DMA	298.15 K	10^7	10	100%	6.5×10^{2}	5.2×10^{2}

S8.2 Sulfuric Acid–Ammonia

Table S7: Particle formation rates obtained from ACDC simulations.

system	T	SA conc. [cm^{-3}]	base conc. [ppt]	RH	$J^{\mathrm{QC+QHA}}$ [cm^{-3}s^{-1}]	$J^{\mathrm{QC+QHA+HL}}$ [cm^{-3}s^{-1}]
SA-AM	278.15 K	10^5	10	0%	$6. \times 10^{-14}$	$9. \times 10^{-13}$
SA-AM	278.15 K	10^5	10	50%	6.5×10^{-14}	1.1×10^{-12}
SA-AM	278.15 K	10^5	10	100%	9.3×10^{-14}	8.7×10^{-13}
SA-AM	278.15 K	10^5	10000	0%	$5. \times 10^{-6}$	1.6×10^{-5}
SA-AM	278.15 K	10^5	10000	50%	5.6×10^{-6}	$2. \times 10^{-5}$
SA-AM	278.15 K	10^5	10000	100%	1.2×10^{-5}	6.6×10^{-5}
SA-AM	278.15 K	10^7	10	0%	6.1×10^{-8}	9.2×10^{-7}
SA-AM	278.15 K	10^7	10	50%	6.6×10^{-8}	1.1×10^{-6}
SA-AM	278.15 K	10^7	10	100%	9.5×10^{-8}	8.8×10^{-7}
SA-AM	278.15 K	10^7	10000	0%	5.	1.4×10^{1}
SA-AM	278.15 K	10^7	10000	50%	5.5	1.8×10^{1}
SA-AM	278.15 K	10^7	10000	100%	1.2×10^{1}	6.1×10^{1}
SA-AM	298.15 K	10^5	10	0%	1.5×10^{-19}	1.1×10^{-18}
SA-AM	298.15 K	10^5	10	50%	1.5×10^{-19}	1.2×10^{-18}
SA-AM	298.15 K	10^5	10	100%	1.8×10^{-19}	7.6×10^{-19}
SA-AM	298.15 K	10^5	10000	0%	$1. \times 10^{-8}$	1.3×10^{-8}
SA-AM	298.15 K	10^5	10000	50%	1.2×10^{-8}	1.5×10^{-8}
SA-AM	298.15 K	10^5	10000	100%	1.7×10^{-8}	5.1×10^{-8}
SA-AM	298.15 K	10^7	10	0%	1.5×10^{-13}	1.2×10^{-12}
SA-AM	298.15 K	10^7	10	50%	1.5×10^{-13}	1.2×10^{-12}
SA-AM	298.15 K	10^7	10	100%	1.8×10^{-13}	7.8×10^{-13}
SA-AM	298.15 K	10^7	10000	0%	$1. \times 10^{-2}$	1.3×10^{-2}
SA-AM	298.15 K	10^7	10000	50%	1.2×10^{-2}	1.5×10^{-2}
SA-AM	298.15 K	10^7	10000	100%	1.7×10^{-2}	$5. \times 10^{-2}$

S8.3 Methanesulfonic Acid–Methylamine

Table S8: Particle formation rates obtained from ACDC simulations.

system	T	SA conc. [cm^{-3}]	base conc. [ppt]	RH	$J^{\mathrm{QC+QHA}}$ [cm^{-3}s^{-1}]	$J^{\mathrm{QC+QHA+HL}}$ [cm^{-3}s^{-1}]
MSA-MA	278.15 K	10^5	1	0%	2.3×10^{-14}	2.6×10^{-14}
MSA-MA	278.15 K	10^5	1	50%	2.4×10^{-14}	3.5×10^{-14}
MSA-MA	278.15 K	10^5	1	100%	5.4×10^{-14}	4.8×10^{-13}
MSA-MA	278.15 K	10^5	100	0%	$8. \times 10^{-7}$	4.2×10^{-7}
MSA-MA	278.15 K	10^5	100	50%	8.4×10^{-7}	4.9×10^{-7}
MSA-MA	278.15 K	10^5	100	100%	1.4×10^{-6}	1.7×10^{-6}
MSA-MA	278.15 K	10^7	1	0%	2.9×10^{-8}	3.3×10^{-8}
MSA-MA	278.15 K	10^7	1	50%	$3. \times 10^{-8}$	4.4×10^{-8}
MSA-MA	278.15 K	10^7	1	100%	6.7×10^{-8}	5.7×10^{-7}
MSA-MA	278.15 K	10^7	100	0%	6.9×10^{-1}	3.8×10^{-1}
MSA-MA	278.15 K	10^7	100	50%	7.4×10^{-1}	4.6×10^{-1}
MSA-MA	278.15 K	10^7	100	100%	1.4	1.7
MSA-MA	298.15 K	10^5	1	0%	2.2×10^{-20}	$1. \times 10^{-20}$
MSA-MA	298.15 K	10^5	1	50%	2.3×10^{-20}	1.2×10^{-20}
MSA-MA	298.15 K	10^5	1	100%	3.8×10^{-20}	1.1×10^{-19}
MSA-MA	298.15 K	10^5	100	0%	2.1×10^{-12}	8.7×10^{-13}
MSA-MA	298.15 K	10^5	100	50%	2.2×10^{-12}	$1. \times 10^{-12}$
MSA-MA	298.15 K	10^5	100	100%	3.6×10^{-12}	5.4×10^{-12}
MSA-MA	298.15 K	10^7	1	0%	2.8×10^{-14}	1.3×10^{-14}
MSA-MA	298.15 K	10^7	1	50%	2.8×10^{-14}	1.5×10^{-14}
MSA-MA	298.15 K	10^7	1	100%	4.9×10^{-14}	1.4×10^{-13}
MSA-MA	298.15 K	10^7	100	0%	2.1×10^{-6}	8.7×10^{-7}
MSA-MA	298.15 K	10^7	100	50%	2.2×10^{-6}	$1. \times 10^{-6}$
MSA-MA	298.15 K	10^7	100	100%	3.6×10^{-6}	5.4×10^{-6}

References

(1) Kubečka, J.; Besel, V.; Neefjes, I.; Knattrup, Y.; Kurtén, T.; Vehkamäki, H.; Elm, J. Computational tools for handling molecular clusters: Configurational sampling, storage, analysis, and machine learning. *ACS Omega* **2023**, *8*, 45115–45128, DOI: 10.1021/acsomega.3c07412.

(2) Zhang, J.; Dolg, M. ABCluster: The artificial bee colony algorithm for cluster global optimization. *Phys. Chem. Chem. Phys.* **2015**, *17*, 24173–24181, DOI: 10.1039/C5CP04060D.

(3) Zhang, J.; Dolg, M. Global optimization of clusters of rigid molecules using the artificial bee colony algorithm. *Phys. Chem. Chem. Phys.* **2016**, *18*, 3003–3010, DOI: 10.1039/C5CP06313B.

(4) McGrath, M. J.; Olenius, T.; Ortega, I. K.; Loukonen, V.; Paasonen, P.; Kurtén, T.; Kulmala, M.; Vehkamäki, H. Atmospheric Cluster Dynamics Code: A flexible method for solution of the birth-death equations. *Atmos. Chem. Phys.* **2012**, *12*, 2345–2355, DOI: 10.5194/acp-12-2345-2012.

(5) Olenius, T.; Kupiainen-Määttä, O.; Ortega, I. K.; Kurtén, T.; Vehkamäki, H. Free energy barrier in the growth of sulfuric acid–ammonia and sulfuric acid–dimethylamine clusters. *J. Chem. Phys.* **2013**, *139*, 084312, DOI: 10.1063/1.4819024.

(6) Elm, J.; Passananti, M.; Kurtén, T.; Vehkamäki, H. Diamines can initiate new particle formation in the atmosphere. *J. Phys. Chem. A* **2017**, *121*, 6155–6164, DOI: 10.1021/acs.jpca.7b05658.

(7) Maso, M. D.; Hyvärinen, A.; Komppula, M.; Tunved, P.; Kerminen, V.; Lihavainen, H.; Öviisanen, Y.; Hansson, H.-C.; Kulmala, M. Annual and interannual

variation in boreal forest aerosol particle number and volume concentration and their connection to particle formation. *Tellus B: Chem. Phys. Meteorol.* **2008**, *60*, 495–508, DOI: 10.1111/j.1600-0889.2008.00366.x.

(8) Kontkanen, J.; Lehtipalo, K.; Ahonen, L.; Kangasluoma, J.; Manninen, H. E.; Hakala, J.; Rose, C.; Sellegri, K.; Xiao, S.; Wang, L. et al. Measurements of sub-3 nm particles using a particle size magnifier in different environments: from clean mountain top to polluted megacities. *Atmos. Chem. Phys.* **2017**, *17*, 2163–2187, DOI: 10.5194/acp-17-2163-2017.

(9) Jensen, A. B.; Elm, J. Massive assessment of the geometries of atmospheric molecular clusters. *J. Chem. Theory Comput.* **2024**, *20*, 8549–8558, DOI: 10.1021/acs.jctc.4c01046.

(10) Coriani, S.; Marchesan, D.; Gauss, J.; Hättig, C.; Helgaker, T.; Jørgensen, P. The accuracy of ab initio molecular geometries for systems containing second-row atoms. *J. Chem. Phys* **2005**, *123*, 184107, DOI: 10.1063/1.2104387.

(11) Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.* **2012**, *18*, 9955–9964, DOI: 10.1002/chem.201200497.

(12) Joel M Bowman, T. C.; Meyer, H.-D. Variational quantum approaches for computing vibrational energies of polyatomic molecules. *Mol. Phys.* **2008**, *106*, 2145–2182, DOI: 10.1080/00268970802258609.

(13) Jacobsen, R. L.; Johnson, R. D. I.; Irikura, K. K.; Kacker, R. N. Anharmonic Vibrational Frequency Calculations Are Not Worthwhile for Small Basis Sets. *J. Chem. Theory Comput.* **2013**, *9*, 951–954, DOI: 10.1021/ct300293a.

(14) Johnson, R. NIST 101. Computational Chemistry Comparison and Benchmark Database. 1999.

(15) Myllys, N.; Elm, J.; Kurtén, T. Density functional theory basis set convergence of sulfuric acid-containing molecular clusters. *Comput. Theor. Chem.* **2016**, *1098*, 1–12, DOI: 10.1016/j.comptc.2016.10.015.

(16) Dunn, M. E.; Evans, T. M.; Kirschner, K. N.; Shields, G. C. Prediction of accurate anharmonic experimental vibrational frequencies for water clusters,$(H_2O)_n$, n = 2–5. *J. Phys. Chem. A.* **2006**, *110*, 303–309, DOI: 10.1021/jp054958y.

(17) Huber, K. P.; Herzberg, G. *Molecular spectra and molecular structure IV. Constants of diatomic molecules*; Springer US: Boston, MA, 1979; pp 8–689, DOI: 10.1007/978-1-4757-0961-2_2.

(18) Shimanouchi, T.; Shimanouchi, T.; of Standards, U. S. N. B. *Tables of Molecular Vibrational Frequencies. Consolidated Volume I*; National Standard Reference Data Series; U.S. Government, 1972.

(19) Otto, K. E.; Xue, Z.; Zielke, P.; Suhm, M. A. The Raman spectrum of isolated water clusters. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9849–9858, DOI: 10.1039/C3CP54272F.

(20) Rognoni, A.; Conte, R.; Ceotto, M. How many water molecules are needed to solvate one? *Chem. Sci.* **2021**, *12*, 2060–2064, DOI: 10.1039/D0SC05785A.

(21) Vogt, E.; Kjaergaard, H. G. Vibrational spectroscopy of the water dimer at jet-cooled and atmospheric temperatures. *Annu. Rev. Phys. Chem.* **2022**, *73*, 209–231, DOI: annurev-physchem-082720-104659.

(22) Hintze, P. E.; Kjaergaard, H. G.; Vaida, V.; Burkholder, J. B. Vibrational and

electronic spectroscopy of sulfuric acid vapor. *J. Phys. Chem. A.* **2003**, *107*, 1112–1118, DOI: 10.1021/jp0263626.

(23) Rozenberg, M.; Loewenschuss, A.; Nielsen, C. J. H-bonded clusters in the trimethylamine/water system: A matrix isolation and computational study. *J. Phys. Chem. A.* **2012**, *116*, 4089–4096, DOI: 10.1021/jp3020035.

(24) Kjaersgaard, A.; Vogt, E.; Hansen, A. S.; Kjaergaard, H. G. Room temperature gas-phase detection and Gibbs energies of water amine bimolecular complex formation. *J. Phys. Chem. A.* **2020**, *124*, 7113–7122, DOI: 10.1021/acs.jpca.0c07399.

(25) Fateley, W.; Miller, F. A. Torsional frequencies in the far infrared—II: Molecules with two or three methyl rotors. *Spectrochim. Acta* **1962**, *18*, 977–993, DOI: 10.1016/0371-1951(62)80104-0.

(26) Li, S.; Kjaergaard, H. G.; Du, L. Infrared spectroscopic probing of dimethylamine clusters in an Ar matrix. *J. Environ. Sci.* **2016**, *40*, 51–59, DOI: https://doi.org/10.1016/j.jes.2015.09.012.

(27) Soulard, P.; Tremblay, B. Vibrational study of methylamine dimer and hydrated methylamine complexes in solid neon supported by ab initio calculations. *J. Mol. Struct.* **2021**, *1236*, 130308, DOI: 10.1016/j.molstruc.2021.130308.

(28) Fernández, L.; Marigliano, A. G.; Varetti, E. The vibrational properties of formic acid as monomer and dimer: A DFT study. *Vib. Spectrosc.* **2005**, *37*, 179–187, DOI: j.vibspec.2004.09.001.

(29) Herzberg, G. *Electronic Spectra and Electronic Structure of Polyatomic Molecules*; Molecular spectra and molecular structure; Van Nostrand, 1966.

(30) Koops, T.; Visser, T.; Smit, W. Measurement and interpretation of the absolute infrared intensities of NH_3 and ND_3. *J. Mol. Struct.* **1983**, *96*, 203–218, DOI: 10.1016/0022-2860(83)90049-2.

(31) Zhang, B.; Yang, S.; Huang, Q.-R.; Jiang, S.; Chen, R.; Yang, X.; Zhang, D. H.; Zhang, Z.; Kuo, J.-L.; Jiang, L. Deconstructing vibrational motions on the potential energy surfaces of hydrogen-bonded complexes. *CCS Chemistry* **2021**, *3*, 829–835.

(32) Telfah, A.; Charifi, Z.; Latelli, N.; Qattan, I. A.; Baaziz, H.; Al-Bataineh, Q. M.; Alsaad, A.; Sabirianov, R. Formation of hydrogen bonding network of methane sulfonic acid at low degree of hydration $(MSA)_m{\cdot}(H_2O)_n$ (m = 1–2 and n = 1–5). *Sci. Rep.* **2024**, *14*, 11252, DOI: 10.1038/s41598-024-61364-0.

(33) Lewandowski, H.; Koglin, E.; Meier, R. J. Computational study of the infrared spectrum of acetic acid, its cyclic dimer, and its methyl ester. *Vib. Spectrosc.* **2005**, *39*, 15–22, DOI: 10.1016/j.vibspec.2004.10.003.

(34) Maroń, M. K.; Shultz, M. J.; Vaida, V. Characterization of the nitric acid–water complex in the infrared and near-infrared region at ambient temperatures in carbon tetrachloride. *Chem. Phys. Lett.* **2009**, *473*, 268–273, DOI: 10.1016/j.cplett.2009.03.071.

(35) McCurdy, P. R.; Hess, W. P.; Xantheas, S. S. Nitric acid–water complexes: Theoretical calculations and comparison to experiment. *J. Phys. Chem. A.* **2002**, *106*, 7628–7635, DOI: 10.1021/jp020257e.

(36) Murphy, W.; Zerbetto, F.; Duncan, J.; McKean, D. Vibrational spectrum and harmonic force field of trimethylamine. *J. Phys. Chem.* **1993**, *97*, 581–595, DOI: 10.1021/j100105a010.