Thermodynamic Benchmarking of Hydrated Atmospheric Clusters in Early Particle Formation

Ivo Neefjes[1], Yosef Knattrup[1], Haide Wu[1], Georg Baadsgaard Trolle[1], Jonas Elm[1], and Jakub Kubečka[1]

[1]Department of Chemistry, Aarhus University, Langelandsgade 140, 8000 Aarhus C, Denmark

Correspondence: Jakub Kubečka (ja-kub-ecka@chem.au.dk)

Abstract. To improve computational modeling of hydrated atmospheric molecular clusters, we systematically evaluated quantum-chemical methods for predicting accurate structural and energetic properties of clusters containing a variety of atmospherically relevant acids and bases, with up to five water molecules. We find that the commonly applied ωB97X-D/6-31++G(d,p) method with DLPNO$^{\text{NormalPNO}}$–CCSD(T$_0$)/aug-cc-pVTZ electronic energy correction is suitable for hydrated clusters. Composite density functional methods such as B97-3c, r^{2}SCAN-3c and ωB97X-3c are effective for pre-screening or modeling large clusters, while the local natural orbital approach LNO–CCSD(T)/aug$'$-cc-pVTZ is well-suited for accurate refinement due to its low memory requirements, high accuracy, and favorable computational scaling. Nevertheless, the ωB97X-3c method has a reasonable accuracy even without the electronic energy correction.

We also assessed thermochemical corrections beyond the conventional harmonic oscillator approximation applied only to the lowest free-energy structure. For the limiting cases of no corrections and the ideal maximum corrections, we calculated hydration distributions and particle formation rates, with a specific emphasis on sulfuric acid–ammonia (SA–AM), sulfuric acid–dimethylamine (SA–DMA), and methanesulfonic acid–methylamine (MSA–MA) clusters. Hydration of small clusters is generally limited, with only selected SA- and MSA-containing clusters showing substantial hydration. Due to the high water concentration in the atmosphere, hydration equilibrates fast, increasing the number of accessible states, and thus stabilizing clusters. However, its effect on cluster formation and new particle formation is highly system dependent.

MSA–MA particle formation rates are more sensitive to hydration than those of SA–AM or SA–DMA, though the enhancement remains modest. Despite being more hydrated than SA–DMA clusters, MSA–MA clusters form new particles at relatively low rates, comparable to SA–AM. Under typical atmospheric conditions, SA–DMA is expected to dominate new particle formation, even at high humidity.

1 Introduction

Aerosol particles—solid and liquid particles suspended in the atmosphere—significantly influence both global climate (Li et al., 2022) and human health (Falcon-Rodriguez et al., 2016; Mei et al., 2018). While some aerosols are emitted directly from sources like sea spray, desert dust, volcanic eruptions, pollen, and fossil fuel combustion, most are formed in the atmosphere through a gas-to-particle conversion process known as new particle formation (NPF) (Kulmala et al., 2013). In NPF, low-volatility gas-phase molecules collide and stick together to form atmospheric molecular clusters. These clusters can continue to

grow into aerosol particles through condensation and coagulation. Aerosol particles impact the climate directly, by scattering incoming solar radiation, and indirectly, by providing a surface onto which water can condense to form clouds (Haywood and Boucher, 2020; Lohmann and Feichter, 2005). The latest IPCC assessment report indicates that aerosol particles are responsible for the largest uncertainty in current climate models (Intergovernmental Panel on Climate Change (IPCC), 2023). This uncertainty is mainly due to limited knowledge about the early stages of NPF, where gas-phase molecules form clusters of $\sim$2 nm in diameter (Tröstl et al., 2016).

While the full range of atmospheric molecules contributing to NPF is still unknown, research has shown that clusters containing various acids and bases can rapidly form under atmospheric conditions. Acid–base clusters are stablized by proton transfer between the acid and base components, forming strongly bound salt. Sulfuric acid (H_2SO_4; SA; Sipilä et al. (2010)) plays a well-established role in NPF, while other acids, such as methanesulfonic acid (CH_3SO_3H; MSA; Dawson et al. (2012)) and nitric acid (HNO_3; NTA; Wang et al. (2020)) have been proposed as potential contributors. Formic acid (HCOOH; FA) and acetic acid (CH_3COOH; ACA), the most common organic acids in the atmosphere (Andreae et al., 1988; Keene et al., 1983; Keene and Galloway, 1984; Galloway et al., 1982; Millet et al., 2015), have been shown in computational studies to enhance NPF (Zhang et al., 2022). The most widely studied bases in atmospheric acid–base clusters are amines, with dimethylamine ($(CH_3)_2NH$; DMA) and trimethylamine ($(CH_3)_3N$; TMA) playing a significant role, while ammonia (NH_3; AM), methylamine (CH_3NH_2; MA), and ethylenediamine ($C_2H_4(NH_2)_2$; EDA) have lower contributions (Almeida et al., 2013; Kurtén et al., 2008; Jen et al., 2016; Myllys et al., 2019; DePalma et al., 2012, 2014; Kirkby et al., 2011).

Water (H_2O; W) is ubiquitous in the atmosphere. At high relative humidities (RH), its concentration can reach $\sim 10^{17}$ cm^{-3}, about 10 orders of magnitude higher than that of particle-forming vapors such as sulfuric acid and bases. While water molecules cannot form pure water clusters on their own under typical atmospheric conditions, they can participate in the formation of clusters with other atmospheric molecules (Carlsson et al., 2020). Several atmospheric measurement studies have investigated the effect of RH on NPF, generally finding an anticorrelation between NPF rates and RH (Birmili and Wiedensohler, 2000; Birmili et al., 2003; Boy and Kulmala, 2002; Laaksonen et al., 2008; Woo et al., 2001). Conversely, controlled laboratory studies indicate that increased RH can positively influence particle formation rates (Duplissy et al., 2016; Merikanto et al., 2016). This discrepancy is believed to result from the fact that, although higher RH can directly boost NPF rates, it can also have indirect effects—like increasing cloud cover—that might reduce NPF in the atmosphere as lowered solar radiation leads to reduced gas-phase oxidation chemistry and the hygroscopic growth of preexisting particles increases the overall condensation sink (CS) factor (Hamed et al., 2011). However, it remains unclear whether water induces a consistent shift in NPF rates or affects them in more complex, condition-dependent ways.

State-of-the-art experimental techniques, such as condensation particle counters and particle size magnifiers, can detect aerosol particles down to sizes of $\sim$1.5–3 nm (McMurry, 2000; Vanhanen et al., 2011). However, these methods provide limited information on the chemical composition of the detected particles. While chemical ionization mass spectrometers (CIMS; Zapadinsky et al. (2019); Passananti et al. (2019); Jokinen et al. (2012)) offer molecular insights into clusters, fragmentation artifacts often distort cluster populations, complicating the characterization of sub-2–3 nm particles. In recent decades, computational chemistry methods have been extensively used to address this challenge (e.g., Vehkamäki et al., 2002; Nadykto

and Yu, 2007; Kurtén et al., 2008; Temelso et al., 2012; Almeida et al., 2013; Liu et al., 2018; Xu et al., 2020; Elm et al., 2020). Numerous studies have focused on pure water clusters, exploring and characterizing their potential energy surface as well as examining their properties such as energetics (e.g., binding energies, HOMO-LUMO gap, and vibrational spectra), geometry of molecular interaction, charge distribution, and dipole moments (Gao et al., 2022; Andersson, 2023; Nguyen et al., 2008; Tribello et al., 2011; García-Argote et al., 2024; Nandi et al., 2021). Accurate modeling of sub-2–3 nm hydrated clusters remains computationally demanding, as they can consist of tens of molecules and include a variety of possible molecular species, necessitating the use of approximations. Hence, in computational studies of atmospheric molecular clusters, water is often excluded to reduce computational costs. This exclusion is based on the assumption that experimental studies are typically conducted under similar relative humidity (RH) conditions, minimizing systematic errors from neglecting water. Nevertheless, several computational studies have specifically investigated hydrated clusters, highlighting the potential role of water in aerosol formation. Ianni and Bandy (2000) combined computational chemistry and classical thermodynamics to examine the hydration distributions of SA monomers and dimers. Kurtén et al. (2007) extended this work to SA–AM clusters, while Henschel et al. (2014, 2016) explored the role of humidity in SA–AM and SA–DMA nucleation. These studies demonstrated that water influences proton transfer in atmospheric acid–base clusters and can either promote or inhibit particle formation rates, depending on the cluster composition and environmental conditions. Similar findings were later reported by Ge et al. (2020), Myllys et al. (2021), and Myllys (2023). With the growing number of potential NPF precursor candidates, multiple studies have investigated the hydration of molecular clusters beyond sulfuric acid systems (Zhu et al., 2014; Xu et al., 2010; Weber et al., 2012, 2014; Miao et al., 2015; Chen et al., 2017; Hu et al., 2017; Zhu et al., 2014; Odbadrakh et al., 2020; Gong et al., 2024; Chen et al., 2020). For instance, Chen et al. (2020) showed that humidity can stabilize MSA–MA clusters, significantly enhancing NPF compared to the dry system. Kildgaard et al. (2018a) developed an advanced method for identifying hydrated cluster geometries, which was later applied to study binding strengths between water and various acids (Kildgaard et al., 2018b; Rasmussen et al., 2020).

A growing body of computational studies on atmospheric clusters increasingly supports the systematic inclusion of water in cluster modeling. To facilitate this integration, we benchmark quantum chemistry methods for their accuracy in describing hydrated clusters. This study specifically focuses on the initial stages of new particle formation, involving freshly nucleated particles a few nanometers in size, whereas hydration effects in subsequent growth stages may differ substantially. We evaluate key properties such as binding electronic energies, cluster geometries, vibrational frequencies, and binding free energies. Furthermore, we analyze the hydration distributions across different cluster sizes and compositions and examine the cluster distribution dynamics of the most relevant systems. Thus, this work not only assesses the accuracy of current methods in describing hydrated clusters but also reveals how explicitly incorporating water can influence conclusions regarding the role of humidity in NPF.

2 Methods

2.1 Molecular system datasets

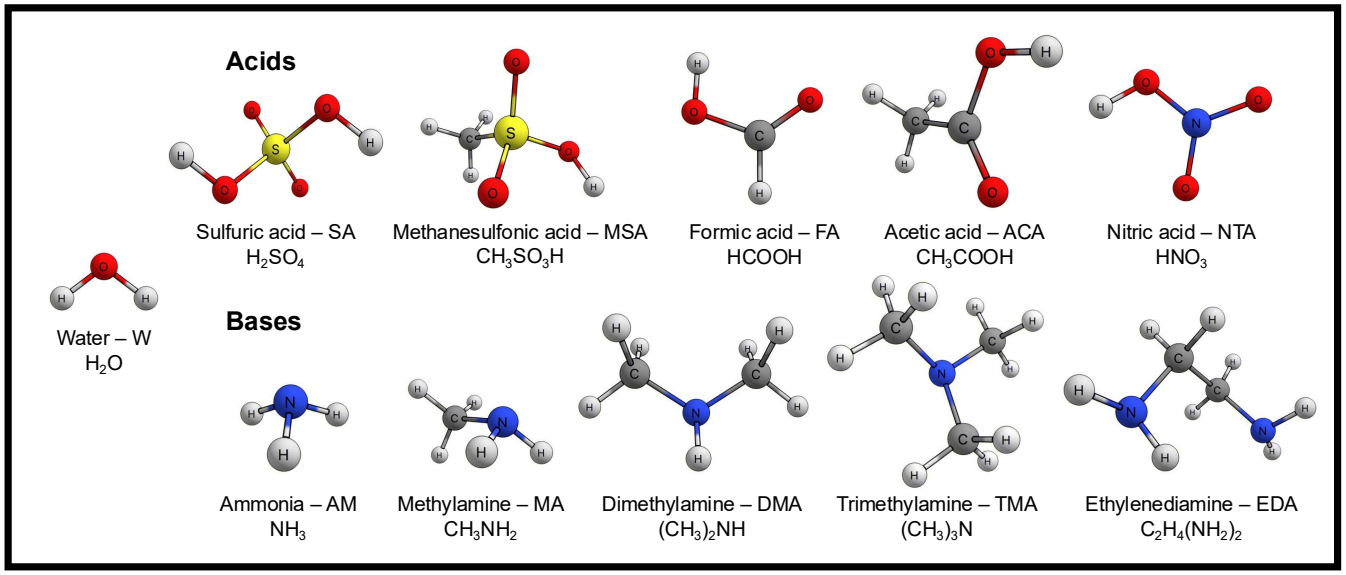

Figure 1. Ball-and-stick representations of the studied monomer molecules.

Microhydrated monomer and dimer clusters formed from various combinations of atmospherically relevant acids and bases, with varying numbers of water molecules, were used to benchmark the accuracy of the quantum chemistry (QC) methods in predicting electronic binding energies, equilibrium geometries, and cluster thermochemistry, and to investigate hydration distributions. The acids and bases included in the study are illustrated in Fig. 1. All combinations that satisfy (acid and/or base)$_{0-2}$W$_{0-5}$ were considered, resulting in a total of 395 unique clusters. For each cluster, we sampled up to five distinct low-energy configurations (< 50 kcal mol^{-1}; see Sec. S1), optimizing their geometries at the GFN1-xTB level of theory. This resulted in a dataset of approximately 1.8k structures.

Microhydrated (sulfuric acid–ammonia)-pair clusters, (SA$_1$AM$_1$)$_{1-6}$W$_{0-10}$, were sampled to investigate how the electronic binding energy error and hydration distribution evolve with cluster size. For each of these cluster compositions, three unique conformers optimized at the GFN1-xTB level of theory were randomly selected from the lowest 50 kcal mol^{-1} configurations to provide a representative sampling of different cluster configurations.

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* were studied using cluster population dynamics to investigate the effect of humidity on the NPF rate.

A detailed description of the configurational sampling (Kubečka et al., 2023; Zhang and Dolg, 2015, 2016) procedure for each dataset is provided in the corresponding sections and Sec. S1.

2.2 Benchmarked quantum chemistry methods

Table 1. Overview of the quantum chemistry methods and basis sets included in this benchmark . Methods are grouped according to whether they use a fixed internal basis set or require a user-selected basis set. For methods requiring a user-selected basis set, one or more of the basis sets listed in the final column were used.

Methods		
Fixed internal basis set	User-selected basis set	**Basis sets**
PM7	M06-2X	6-31+G(d)
GFN1-xTB	PW91	6-31++G(d,p)
GFN2-xTB	ωB97X-D	6-311++G(d,p)
AMC-xTB	RI-MP2	6-311++G(3df,3pd)
GFN1repar	DLPNO–CCSD(T_0)	(aug-)cc-pVDZ
B97-3c	DLPNO–CCSD(T_0)-F12	(aug-)cc-pVTZ
r^2SCAN-3c	LNO–CCSD(T)	aug-cc-pVQZ
ωB97X-3c	CCSD(T*)-F12	

We benchmarked a range of QC methods—from semi-empirical to high-accuracy wavefunction-based approaches—for their accuracy in predicting electronic binding energies and equilibrium geometries of atmospherically relevant hydrated clusters (Tab. 1). PM7 (Stewart, 2012) is a semi-empirical method based on the Hartree–Fock (HF) formalism. GFN1-xTB (Grimme et al., 2017) and GFN2-xTB (Bannwarth et al., 2019), developed by the Grimme group, are density-functional tight-binding methods. AMC-xTB (Knattrup et al., 2024) and GFN1repar (Wu et al., 2024) are reparameterizations of GFN1-xTB tailored for calculations of atmospheric molecular cluster equilibrium structures and electronic binding energies.

We also included empirically corrected DFT methods (DFT-3c) such as B97-3c (Brandenburg et al., 2018), r^2SCAN-3c (Grimme et al., 2021), and ωB97X-3c (Müller et al., 2023), which enhance accuracy in intermolecular interactions through systematic error cancellation while maintaining computational efficiency. Additionally, we assessed hybrid and meta-generalized gradient approximation (GGA) functionals like ωB97X-D (Chai and Head-Gordon, 2008) and M06-2X (Zhao and Truhlar, 2007), along with the GGA functional PW91 (Burke et al., 1998). These functionals have demonstrated reliable thermochemistry and relative binding energies for dry molecular clusters, with ωB97X-D particularly noted for its consistently accurate performance (Elm and Mikkelsen, 2014; Schmitz and Elm, 2020; Jensen et al., 2022).

Last, we included the more computationally intensive wavefunction-based method RI-MP2 (Weigend et al., 1998), along with the domain-based local pair natural orbital (DLPNO; (Riplinger and Neese, 2013) and local natural orbital (LNO; Rolik et al. (2013); Nagy et al. (2018); Nagy and Kállay (2019); Kállay et al. (2020, 2025) coupled cluster methods with single, double, and perturbative triple excitations (CCSD(T)), providing robust electron correlation treatments suitable for high-precision calculations, albeit at a high computational cost.

Together, this selection offers a comprehensive range of methods, covering various levels of accuracy and computational efficiency.

We utilized several basis sets for the QC methods that require these to be explicitly set. For M06-2X and PW91, we employed the Pople basis sets 6-31+G(d), 6-31++G(d,p), and 6-311++G(d,p) (Clark et al., 1983; Ditchfield et al., 1971; Francl et al., 1982; Gordon et al., 1982; Hariharan and Pople, 1973; Hehre et al., 1972; Spitznagel et al., 1987; Krishnan et al., 1980; McLean and Chandler, 1980). For RI-MP2, DLPNO–CCSD(T_0), and LNO–CCSD(T), we used the augmented correlation-consistent basis sets aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ (Dunning, 1989; Kendall et al., 1992; Woon and Dunning, 1993). The Pople basis sets are typically more efficient, while the augmented correlation-consistent basis sets offer higher accuracy (Pitman et al., 2023). We evaluated the performance of ωB97X-D using all the aforementioned basis sets, as well as the 6-311++G(3df,3pd) basis set. For LNO–CCSD(T), we employed the aug′-cc-pVTZ basis set, where the diffuse functions on hydrogen atoms are removed. This variant is often used in noncovalent interaction and cluster studies, as diffuse functions on hydrogen typically contribute little to accuracy but can significantly increase computational cost and cause convergence issues, particularly in correlated wavefunction calculations (Del Bene, 1993). Last, the explicitly correlated (F12) technique was employed for the CCSD(T*)-F12 and DLPNO–CCSD(T_0)-F12 methods (Pavošević et al., 2017) with the cc-pVDZ-F12 and cc-pVTZ-F12 basis set (Peterson et al., 2008).

The DLPNO–CCSD(T_0) and LNO–CCSD(T) methods enhance efficiency compared to traditional coupled cluster methods by using a truncated set of localized or electron pair-specific natural orbitals. The choice of truncation criteria affects the number of natural orbitals included in the calculations. Tighter criteria incorporate more orbitals, improving accuracy but increasing computational cost. We evaluated the NormalPNO, TightPNO, HFC1, and HFC2 settings for DLPNO–CCSD(T_0) and the Normal and Tight settings for LNO–CCSD(T).

QC calculations were performed with the xtb 6.7.0 (Bannwarth et al., 2021), Gaussian16 Rev.B.01 (Frisch et al., 2016), MRCC (Kállay et al., 2020, 2025), and ORCA 5.0.4 and 6.0.1 (Neese, 2012, 2022) programs. This study coincided with the release of ORCA 6.0.1, and, as a test, the B97-3c and r^{2}SCAN-3c methods were recalculated using both versions. While the differences in calculated binding energies were negligible ($\Delta < 0.002$ kcal mol^{-1}), we observed an average decrease in computation time of approximately 10% for the newer version.

2.3 Electronic binding energy benchmark

Using both the *microhydrated monomer and dimer clusters* and *(sulfuric acid–ammonia)-pair clusters*, with geometries optimized at the GFN1-xTB level of theory (see Sec. 2.1), all QC methods were benchmarked based on their electronic binding energy ΔE_{el} for the given geometries:

$$\Delta E_{\mathrm{el}} = E_{\mathrm{el,cluster}} - \sum_i E_{\mathrm{el},i}, \tag{1}$$

where E_{el} is the electronic energy of the cluster/monomer, and the summation runs over all acid, base, and water molecules in the cluster.

The quality of the benchmarked methods was assessed using signed/absolute errors and the mean absolute error (MAE), across different configurations of a cluster compared to a reference method (REF), with

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\Delta E_{\mathrm{el},i}^{\mathrm{REF}} - \Delta E_{\mathrm{el},i}|, \tag{2}$$

where n is the number of configurations, and $\Delta E_{\mathrm{el},i}^{\mathrm{REF}}$ and $\Delta E_{\mathrm{el},i}$ denote the electronic binding energies obtained with the reference method and the benchmarked QC method, respectively.

The CCSD(T*)-F12/cc-pVTZ-F12 method was used as a reference. This CCSD(T) method was chosen because it is widely regarded as the golden standard for calculating energetics (Ramabhadran and Raghavachari, 2013; Kodrycka and Patkowski, 2019), while still being computationally feasible for the small cluster sizes considered here. Kruse et al. (2020) showed that the MAE of this basis set compared to the complete basis set (CBS) limit is 0.04 kcal mol^{-1} when tested on various molecular dimers (S66 dataset; Řezáč et al. (2011)). Schmitz and Elm (2020) reported similar errors (in their SI) from CBS extrapolation for atmospheric acid–base dimers. Rescaled to our systems, we expect the reference method to have a maximum basis set incompleteness error (BSIE) of 0.1 kcal mol^{-1}. Because CCSD(T*)-F12/cc-pVTZ-F12 is computationally prohibitive for all but the smallest clusters, we compared other methods to it and used the best-performing method in terms of efficiency and accuracy, DLPNO$^{\mathrm{NormalPNO}}$-CCSD(T_0)/aug-cc-pVTZ (see Secs. 3.1 and 3.2), as the reference for larger clusters.

All QC output is stored in the Atmospheric Cluster Database (ACDB; Elm and Kubečka (2024); Elm (2019); Kubečka et al. (2023); see Sec. S2).

2.4 Equilibrium geometry benchmark

To evaluate how well different efficient QC methods approximate equilibrium geometries of hydrated clusters and to assess the correlations between them, we compared up to five configurations across methods for the (acid and/or base)$_{0–2}$W$_{0–5}$ clusters. All sampled (GFN1-xTB-equilibrium) geometries successfully reoptimized at the ωB97X-D/6-31++G(d,p) level of theory (see Sec. S1) were further reoptimized with each benchmarked method. The ArbAlign program was used to align the optimized geometries and calculate root-mean-square deviations (RMSD) between them (Temelso et al., 2017). The methods were compared with each other to identify inter-method correlations and against the RI-MP2/aug-cc-pVQZ reference, which is known to provide accurate geometries compared to higher levels of theory (e.g., DF-CCSD(T*)-F12b/cc-pVDZ-F12) (Jensen and Elm, 2024; Coriani et al., 2005).

2.5 Thermochemistry benchmark

QC combined with statistical thermodynamics enables the calculation of thermochemical properties of molecular clusters, such as their Gibbs free energies of formation. The accuracy of such data is difficult to assess due to little experimental data being available. In Sec. 3.3, we place particular emphasis on vibrational frequencies, which are used to construct the vibrational partition function for Gibbs free energy calculations. The ability of a QC method to produce accurate vibrational frequencies

therefore serves as an indicator of its reliability in predicting thermochemical data. In addition, we investigate other potential sources of error in computed thermochemical properties and evaluate their possible magnitudes.

2.6 Hydration distributions

We examined hydration distributions for the (acid and/or base)$_{0–2}$W$_{0–5}$ systems, as well as hydrated clusters of SA–AM, at 278.15 K and 298.15 K. The QC methods that performed best in terms of efficiency and accuracy, based on our benchmarks, were used for these calculations. The population x_n of a cluster containing n water molecules is given by (Henschel et al., 2014)

$$x_n = \left(\frac{p_{\mathrm{H_2O}}}{p^0}\right)^n x_0 \cdot e^{-\Delta_{\mathrm{hydr}}G_n/k_{\mathrm{B}}T}, \tag{3}$$

where the population of the dry cluster, x_0, is chosen such that $\sum_{n=1} x_n = 1$, with the sum extending to the most hydrated cluster considered. Here, $p_{\mathrm{H_2O}}$ is the water partial pressure, p^0 the reference pressure (1 atm), and $\Delta_{\mathrm{hydr}}G_n$ the standard Gibbs free energy of hydration, i.e. $\Delta_{\mathrm{hydr}}G_n = \Delta G_n - \Delta G_0$. Relative humidity (RH) is calculated with respect to saturation vapor pressure ($p^0_{\mathrm{H_2O}}$) as RH = $p_{\mathrm{H_2O}}/p^0_{\mathrm{H_2O}} \cdot 100\%$, while we obtained $p^0_{\mathrm{H_2O}}$ using the August–Roche–Magnus equation (Alduchov and Eskridge, 1996; Westermann et al., 2016).

2.7 Particle formation rate calculations

To investigate the impact of hydration on NPF, we calculated particle formation rates (J) for the SA$_{0-3}$AM$_{0-3}$W$_{0-5}$, SA$_{0-3}$DMA$_{0-3}$W$_{0-5}$, and MSA$_{0-3}$MA$_{0-3}$W$_{0-5}$ systems using the Atmospheric Cluster Dynamics Code (ACDC) (McGrath et al., 2012; Olenius et al., 2013; Olenius, 2018). Cluster evaporation rates were derived from the binding Gibbs free energies of the clusters, calculated using the most reliable methods identified in the electronic energy and equilibrium geometry benchmarks.

J was evaluated over a 0–100% relative humidity range at temperatures of 278.15 K and 298.15 K. The following constant monomer concentrations were used: SA and MSA at 10^5 and 10^7 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.

Coagulation loss (CL) of clusters was included using CL $= 10^{-3}(d/d_{\mathrm{SA}})^{-1.6}$ s^{-1}, where d is the cluster diameter and d_{SA} that of the SA monomer (Maso et al., 2008). Clusters larger than (acid)$_3$(base)$_3$ were considered as spontaneously outgrowing into particles. For instance, acid, base, and cluster addition to the hydrated (acid)$_3$(base)$_3$ clusters were allowed to grow out of the simulations and contribute to the particle formation rate. A more detailed description of the ACDC simulations is provided in Sec. S3.

3 Results

3.1 Electronic binding energy benchmark

3.1.1 Hydrated monomers and dimers

Single point calculations at the CCSD(T*)-F12/cc-pVTZ-F12 reference method were feasible for only ∼400 small clusters within a reasonable computation time on our hardware. To assess alternative reference methods, we first compared high-quality RI-MP2, DLPNO, and LNO approaches against CCSD(T*)-F12/cc-pVTZ-F12 for this subset. The top part of Fig. 2 presents violin plots of the absolute errors alongside the average CPU time per single-point energy calculation. Calculations were run on either Intel Xeon Gold 6248R or Intel Xeon Platinum 8358 CPUs. Since CPU time is hardware-dependent, the reported times are only indicative. Several tested methods closely agree with CCSD(T*)-F12/cc-pVTZ-F12. Among them, DLPNO$^{\text{NormalPNO}}$-CCSD(T_0)/aug-cc-pVTZ stands out with a low mean absolute error (MAE) of ∼0.20 kcal mol^{-1}, low memory requirements, and an average computational cost under 1 CPU hour. This is consistent with the findings of Schmitz and Elm (2020). We therefore selected it as the reference method for the full dataset of ∼1.8k structures. However, we also highlight the accuracy of the LNO methods, along with their low memory requirements, which stem from the use of local MP2 natural orbitals. In contrast, DLPNO requires computing and storing pair natural orbitals. For example, single-point calculations on $SA_1TMA_1W_5$ demonstrate that LNO$^{\text{Tight}}$-CCSD(T)/aug-cc-pVTZ requires only 2 GB of memory, whereas DLPNO$^{\text{NormalPNO}}$-CCSD(T_0)/aug-cc-pVTZ requires 6 GB. We chose the DLPNO method over the LNO methods for consistency with previous studies rather than for their accuracy difference.

In the full dataset comparison, the semi-empirical methods are extremely fast, with mean CPU times of 0.4–3 seconds. Among them, GFN1-xTB and GFN2-xTB exhibit the lowest MAEs, with GFN1-xTB showing a slightly lower maximum in absolute electronic binding energy error. For many dry clusters, GFN1-xTB has been found to outperform GFN2-xTB as well (Jensen et al., 2022; Rasmussen et al., 2022; Wu et al., 2023; Engsvang and Elm, 2022). However, GFN2-xTB is significantly faster, requiring only 0.5 seconds per calculation compared to 3 seconds for GFN1-xTB. AMC-xTB, a reparameterization of GFN1-xTB for dry molecular clusters, performs worse than the original GFN1-xTB for these microhydrated monomers and dimers, likely because no water-containing clusters were included during the reparameterization.

The r^{2}SCAN-3c and ωB97X-3c methods perform particularly well, with MAEs of 1.29 and 1.26 kcal mol^{-1} but still with maximum absolute errors of ∼6.5 and ∼6.1 kcal mol^{-1}, respectively. However, ωB97X-3c requires more than three times the CPU time of the other two DFT-3c methods studied.

In Sec. S4, we show signed electronic binding energy errors for all benchmarked QC methods. Among the M06-2X, PW91, and ωB97X-D functionals, ωB97X-D performs best when using the same basis set. Paired with the aug-cc-pVQZ basis set, ωB97X-D achieves an MAE of 1.25 kcal mol^{-1}, comparable to the 1.26 kcal mol^{-1} of ωB97X-3c, though at a significantly higher computational cost of 35 CPU hours compared to just 7 CPU minutes for ωB97X-3c. In contrast, the commonly used combination of ωB97X-D with the 6-31++G(d,p) basis set has a similar CPU time of 7 CPU minutes to ωB97X-3c but results

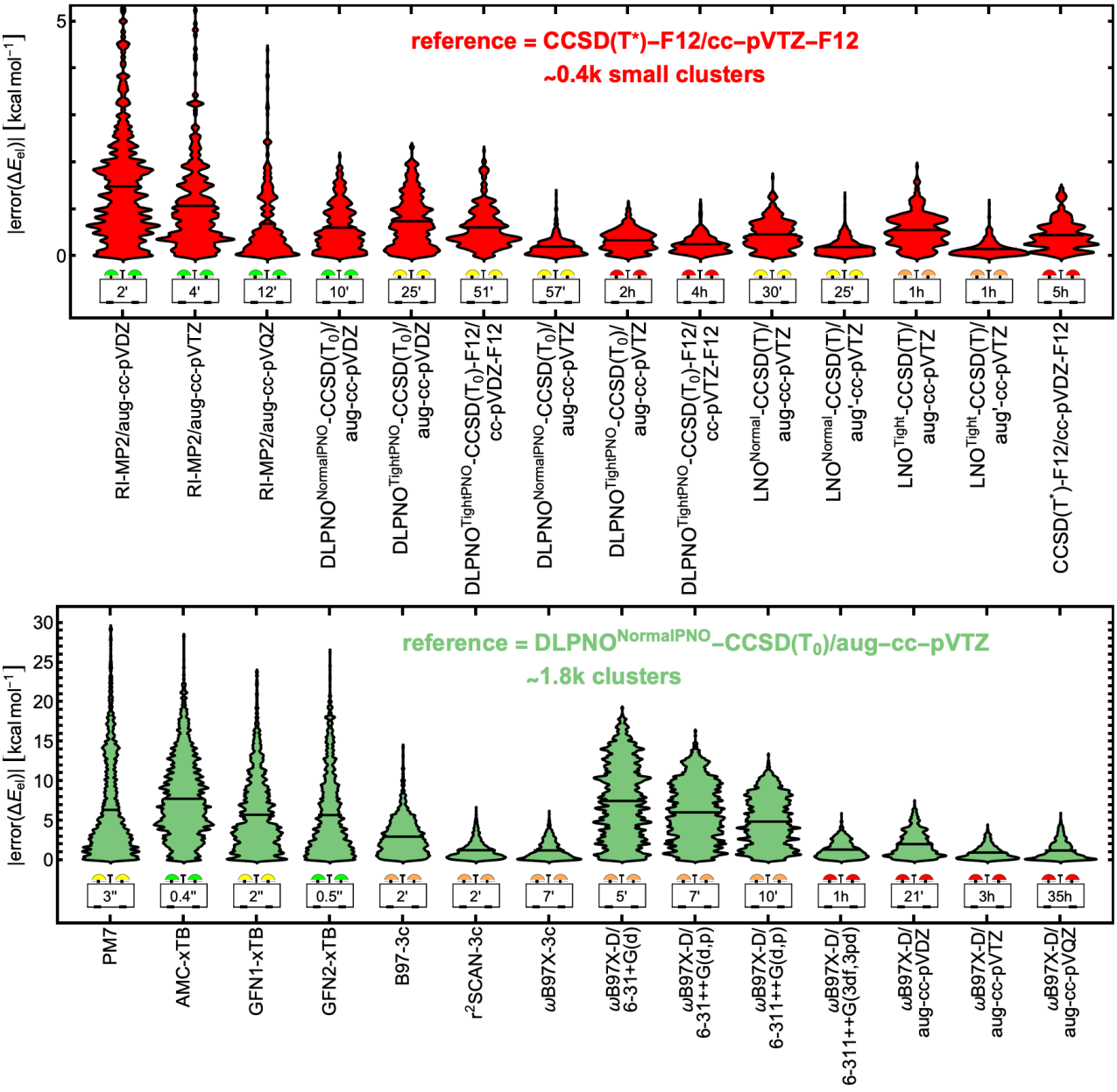

Figure 2. Violin plots of absolute errors in electronic binding energies for microhydrated monomers and dimers. Height indicates error magnitude, width represents the number of configurations with the same error, and vertical black lines mark mean absolute errors. The dataset includes all 395 unique combinations of the molecules in Fig. 1, with formula (acid and/or base)$_{0–2}$(water)$_{0–5}$. Top: RI-MP2, DLPNO, and LNO methods, benchmarked against CCSD(T*)-F12/cc-pVTZ-F12 for a subset of ∼0.4k small clusters. Bottom: Semi-empirical, DFT-3c, and ωB97X-D methods compared to DLPNO$^{\text{NormalPNO}}$-CCSD(T$_0$)/aug-cc-pVTZ for the full data set of ∼1.8k clusters. Average CPU times are given in seconds (″), minutes (′), or hours (h).

in a significantly higher MAE of 6.07 kcal mol^{-1}. For comparison, the DLPNO$^{\text{NormalPNO}}$-CCSD(T$_0$)/aug-cc-pVTZ reference method required an average of 5.3 CPU hours for the full dataset.

Based on this electronic binding energy benchmark, the r^{2}SCAN-3c and ωB97X-3c methods stand out as excellent choices for fast and accurate binding energy calculations of hydrated acid–base clusters.

3.1.2 Scaling with cluster size

To examine how the accuracy of electronic binding energies from the QC methods evolves with increasing cluster size, we analyzed all clusters satisfying the composition $(SA_1AM_1)_{1-6}W_{0-10}$. Only a subset of the best-performing QC methods—those offering a balance of efficiency and accuracy based on the electronic binding energy benchmark in the previous section—was included. The DLPNO$^{\text{NormalPNO}}$-CCSD(T$_0$)/aug-cc-pVTZ method was again used as the reference, as it remains computationally feasible even for relatively large clusters.

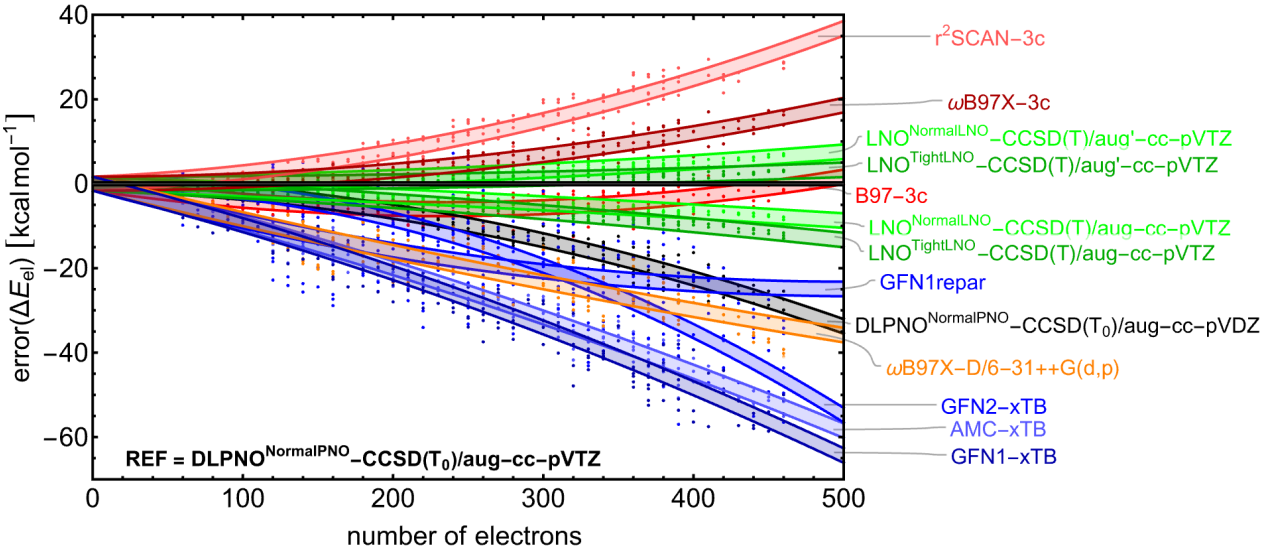

Figure 3. The signed error in (electronic) binding energy (kcal mol^{-1}) relative to the DLPNO$^{\text{NormalPNO}}$-CCSD(T$_0$)/aug-cc-pVTZ reference calculations as a function of the number of electrons for multiple quantum chemistry methods. Electronic energies were determined from single-point energy calculations for three conformers across all clusters with the composition $(SA_1AM_1)_{1-6}W_{0-10}$, randomly selected from the GFN1-xTB optimized geometries. A second-order polynomial, constrained to pass through the origin [0,0], was fit to the data for each quantum chemistry method. The shaded area illustrates the 90% confidence intervals for individual predictions based on this fit.

Figure 3 shows the signed error in electronic binding energy as a function of the number of electrons in the system for each of the included QC methods relative to the reference. For visualization, a second-order polynomial passing through the origin was fitted to the data for each QC method, with the shaded area representing the 90% confidence interval for single predictions. As expected, the semi-empirical xTB methods show the largest errors, with values reaching -60 kcal mol^{-1} for the largest clusters. AMC-xTB, a reparameterization of GFN1-xTB tailored to reproduce ωB97X-D/6-31++G(d,p) data

for dry atmospheric clusters, performs slightly better than GFN1-xTB. For this benchmark, we also included the GFN1repar method, a reparameterization of GFN1-xTB aimed at reproducing B97-3c data for large dry atmospheric clusters (fitted up to $(acid)_{10}(base)_{10}$ clusters). GFN1repar shows significant improvement in accuracy as cluster size increases, outperforming the other semi-empirical methods. While the other semi-empirical methods exhibit a more than linear increase in error magnitude with cluster size, GFN1repar demonstrates a slightly less-than-linear trend, which can be attributed to its focus on large clusters and its fitting to B97-3c, a method that performs better but exhibits similar behavior.

The performance of DLPNO$^{\text{NormalPNO}}$-CCSD(T_0) with a double zeta basis set highlights the significance of basis set size, as an error of approximately -30 kcal mol^{-1} in electronic binding energy is observed for the largest clusters studied here. The low errors associated with LNO–CCSD(T) methods (even lower with tighter LNO criteria and reduced diffuse functions for hydrogens) validate the choice of the reference method used in this study. Given its high memory efficiency compared to DLPNO, LNO methods are recommended for calculating the properties of large clusters in future studies (Knattrup and Elm, 2025).

The DFT-3c methods B97-3c and ωB97X-3c perform exceptionally well, especially considering their efficiency. Although the error of ωB97X-3c increases with the number of electrons, it does so less steeply than r^{2}SCAN-3c and ωB97X-D/6-31++G(d,p), resulting in absolute errors lower than 20 kcal mol^{-1} for the largest cluster studied here. B97-3c shows an increasingly negative error up to $\sim$250 electrons, after which the error magnitude decreases, resulting in a small positive error around 500 electrons. The absolute errors stay below 10 kcal mol^{-1} for all studied cluster sizes, with a remarkably low error for the largest cluster due to the aforementioned trend.

For the SA–AM–W clusters studied here, the DFT-3c methods outperform all other fast methods regarding the accuracy of electronic binding energies relative to the chosen reference method. Given its exceptional performance in both the cluster size benchmark and the binding energy analysis of microhydrated monomers and dimers in the previous section, along with being approximately 1 to 2 orders of magnitude faster than the reference, ωB97X-3c could serve as an efficient method for obtaining reasonable electronic binding energies for large hydrated atmospheric clusters. This conclusion aligns with recent findings for small dry clusters reported by Jensen and Elm (2024).

3.2 Equilibrium Cluster Geometry

Figure 4 shows the mean RMSD for selected methods, while results for the other methods are provided in Sec. S5. All $\sim$1.8k optimization were performed with default setting, except for the RI-MP2/aug-cc-pVQZ reference, where only $\sim$0.6k converged with extreme SCF and very tight optimization criteria. Compared to the reference, both ωB97X-D/6-31++G(d,p) and ωB97X-3c perform best with an RSMD of 0.04 Å. Besides the DFT-3c methods, ωB97X-D/6-31++G(d,p) clearly performs the best, as has also been shown in previous studies (Jensen and Elm, 2024). However, ωB97X-3c yields a similar accuracy, alongside accurate electronic binding energy benchmarks presented in Section 3.1. r^{2}SCAN-3c also shows good agreement with the RI-MP2/aug-cc-pVQZ reference. Given that r^{2}SCAN-3c is significantly more computationally efficient than ωB97X-D/6-31++G(d,p) and ωB97X-3c, it is well-suited for use as an intermediate optimization method during configurational sampling of hydrated clusters. Interestingly, GFN2-xTB performs better than GFN1-xTB. The opposite conclusion was reported for dry

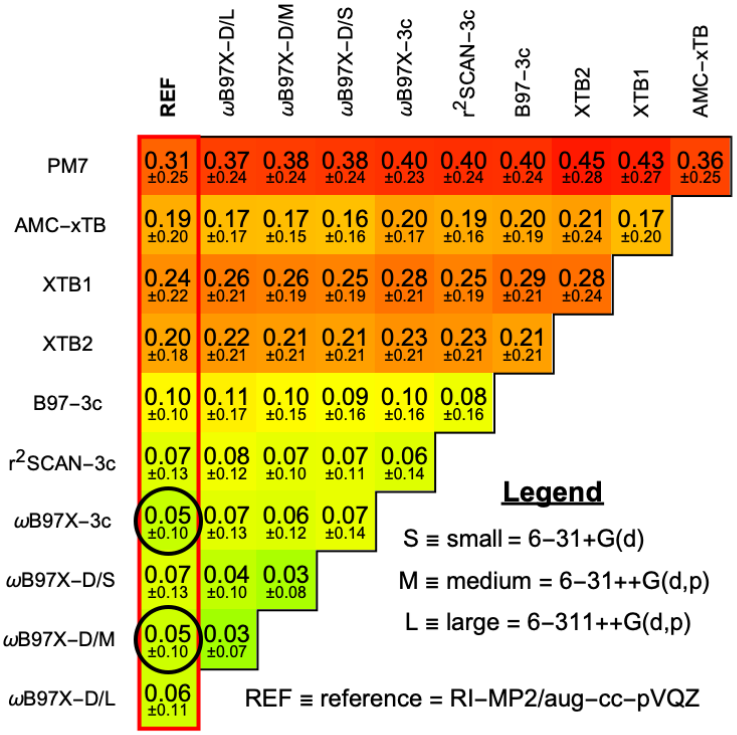

Figure 4. Mean of root-mean-square deviations (RMSD [Å]) between (acid and/or base)$_{0–2}$W$_{0–5}$ equilibrium geometries optimized at different levels of theory. Black circles indicate the best-performing methods relative to the reference, RI-MP2/aug-cc-pVQZ (REF). More detailed comparison is presented in Sec. S5.

atmospheric clusters (Jensen and Elm, 2024). While less accurate, the xTB methods are fast and thus suitable for geometry pre-optimization. The PM7 method seems less suitable for molecular clusters.

3.3 Cluster Thermochemistry

The thermodynamics of molecular systems arise from their dynamics on the potential energy surface (PES), where they transition between discrete vibrational–rotational–translational energy levels associated with different configurational minima. Although the previous sections indicated a low MAE across QC methods for electronic binding energies and equilibrium geometries at the GFN1-xTB level—suggesting that the PES shape is generally well reproduced—this does not necessarily guarantee that thermochemical properties are accurately captured.

Benchmarking thermochemical properties of atmospheric molecular clusters is particularly challenging due to the scarcity of reliable reference data, and is therefore often omitted in methodological evaluations. In this section, we explicitly address thermochemical aspects of hydrated molecular clusters. The dominant contributions arise from vibrations around the most pop-

ulated minimum-energy conformer. To improve accuracy beyond the harmonic oscillator approximation, we consider standard corrections, including anharmonicity scaling of vibrational frequencies, quasi-harmonic treatment of low-frequency modes, and multi-conformer Boltzmann averaging to account for transitions between multiple low-energy conformers. For comparison, we will also derive thermochemical corrections from umbrella sampling, as demonstrated by Kubečka et al. (2025).

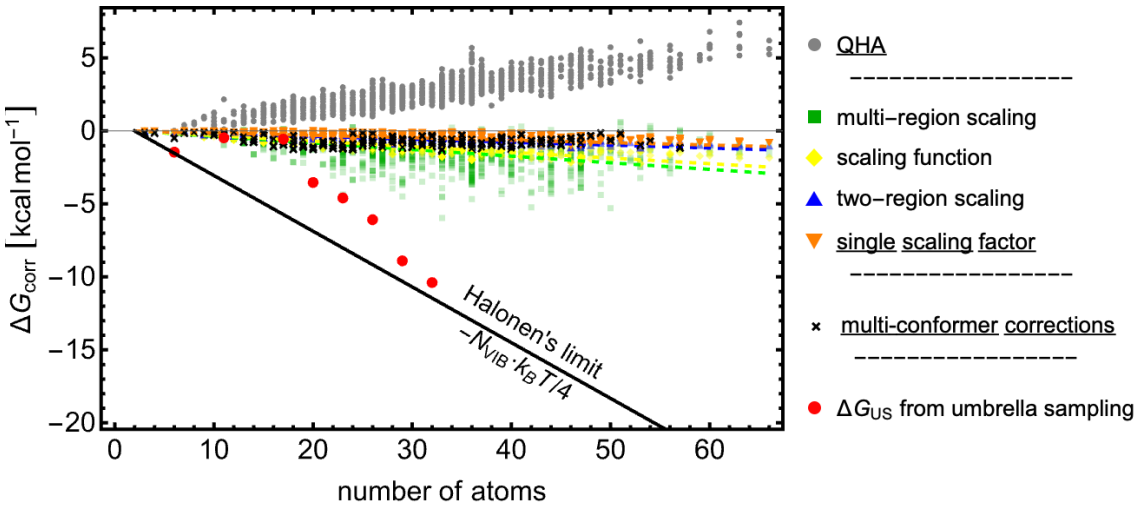

Figure 5. Free energy contributions from vibrational corrections (colored points corresponding to different scalings) and from accounting for multiple low-lying free energy minima, calculated using Eq. 4 as derived by Partanen et al. (2016) (black circles).

Figure 5 shows the free energy contributions from these corrections as a function of the number of atoms in the cluster. Halonen (2024) used MD simulations with force-field methods to demonstrate that the combined effects of anharmonicity and interconversion between minima can reach up to $k_{\mathrm{B}}T/4$ per vibrational mode. We refer to this upper bound as the Halonen thermodynamic limit (black line).

Each correction is discussed in detail in the following subsections. While a full quantitative treatment of anharmonicity is beyond the scope of this study, discussing these corrections and estimating their expected magnitudes provides a more physically realistic description of the vibrational and conformational contributions to the free energies of hydrated molecular clusters.

3.3.1 Low-vibrational frequency treatment

Low-vibrational frequencies are common in hydrated molecular clusters and may even play an important role in their stabilization. However, within the harmonic oscillator treatment and due to numerical inaccuracies, such frequencies can be underestimated, leading to too low values. This poses a problem because the entropic contribution diverges to $-\infty$ as the vibrational frequency approaches zero. To mitigate this issue, the quasi-harmonic approximation (QHA; (Grimme, 2012)) replaces the vibrational entropy of low-frequency modes with the corresponding rotational entropy. In this work, we applied a smooth rotor–vibration transition with crossover frequency of 100 cm^{-1}. As shown in Fig. 5, the QHA correction to the Gibbs

free energy is generally positive and increases approximately linearly with the number of atoms, reaching $\sim$5 kcal mol^{-1} for a 60-atom cluster. It is worth noting that some programs (e.g., ORCA and XTB) already apply this approximation by default.

3.3.2 Vibrational anharmonicity correction

Vibrational frequencies are typically calculated within the rigid-rotor harmonic oscillator (RRHO) approximation, which neglects anharmonicity. This harmonic treatment is a significant source of error when comparing calculated vibrational frequencies to experimental results. Halonen (2024) highlighted that anharmonicity becomes increasingly important for larger clusters, as the number of vibrational modes grows with cluster size. The harmonic approximation generally overestimates vibrational frequencies by 2–6% (Lin et al., 2008).

Second-order vibrational perturbation theory (VPT2; Joel M Bowman and Meyer (2008)), as implemented in ORCA (Barone et al., 2014), can be used to account for anharmonicity. VPT2 typically reproduces experimental fundamental vibration frequencies to within 0–30 cm^{-1}. However, due to practical limitations such as computational cost and convergence issues, vibrational scaling factors are often applied as a simpler alternative. While scaling factors are commonly derived empirically to improve agreement with experiment and thus correct several sources of systematic error (e.g., missing anharmonicity, basis-set incompleteness, and method deficiencies), they are also frequently used as a practical approximation to anharmonic corrections when explicit anharmonic calculations are not feasible. In this work, we follow the latter strategy by deriving scaling factors from comparisons between harmonic and VPT2 anharmonic frequencies, consistent with established practice in cluster studies (e.g., Temelso et al., 2011). Jacobsen et al. (2013) further note that, for the small basis sets and methods they tested, scaled anharmonic vibrational frequencies were not significantly more accurate than scaled harmonic ones when compared with experiment, indicating that explicit anharmonic calculations provided limited additional accuracy in that regime.

Scaling factors have been defined for various methods (Johnson, 1999; Myllys et al., 2016; Tikhonov et al., 2024). However, we fitted our own scaling corrections for the (acid and/or base)$_{0\text{–}2}$W$_{0\text{–}5}$ clusters, testing four approaches: (1) a single constant scaling factor across the full 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}, and (4) a flexible scaling function of the form $A - B \cdot \nu_{\mathrm{harm}} - C/(D + \nu_{\mathrm{harm}})$. The details of the fitting and analysis are provided in Sec. S6. Here, we only summarize that we find that a single scaling factor is sufficient, with no significant improvement from more complex corrections. Table 2 shows the MAEs relative to experimental (Dunn et al., 2006; Huber and Herzberg, 1979; Shimanouchi et al., 1972; Otto et al., 2014; Rognoni et al., 2021; Vogt and Kjaergaard, 2022; Hintze et al., 2003; Rozenberg et al., 2012; Kjaersgaard et al., 2020; Soulard and Tremblay, 2021; Fateley and Miller, 1962; Li et al., 2016; Fernández et al., 2005; Herzberg, 1966; Koops et al., 1983; Zhang et al., 2021; Telfah et al., 2024; Lewandowski et al., 2005; Maroń et al., 2009; McCurdy et al., 2002) vibrational frequencies with and without applying a single scaling factor. Ideally, scaling factors would be fitted directly to experimental fundamentals, but the lack of consistent, unambiguous vibrational assignments for many of the studied clusters makes this infeasible; we therefore adopt the internally consistent and fully automatable harmonic–VPT2 approach.

The scaling factor of 0.950 for r2SCAN-3c is slightly lower than the absolute scaling factor of 0.9688 reported by Tikhonov et al. (2024), a difference that lies within the expected variation when scale factors are fitted to different benchmarking sets. For reference,

Table 2. The single scaling factor obtained by fitting to the difference between the anharmonic (VPT2) and harmonic vibrational frequencies compared to list of experimental observations. See details in Sec. S6.

	B97-3c	r^{2}SCAN-3c	ωB97X-3c	ωB97X-D/6-31++G(d,p)	RI-MP2/aug-cc-pVQZ
no scaling (scaling factor)	<1>	<1>	<1>	<1>	<1>
MAE [cm^{-1}]	61	86	131	110	97
single scaling factor	<0.944>	<0.950>	<0.954>	<0.950>	<0.95 (Johnson, 1999)>
MAE [cm^{-1}]	81	52	60	**39**	40

the two factors differ by more than our MAE of 52 cm^{-1} only for harmonic frequencies above roughly 2766 cm^{-1}. It is also worth noting that our two-region scaling (see Sec. S6.4) yields a factor of 0.969 for modes below 2000 cm^{-1}, essentially identical to the 0.9688 reported by Tikhonov et al. (2024).

For r^{2}SCAN-3c, ωB97X-3c, ωB97X-D/6-31++G(d,p), and RI-MP2/aug-cc-pVQZ scaling significantly improves the MAE compared to the unscaled harmonic approximation. In contrast, B97-3c yields the lowest MAE without scaling, and applying scaling functions actually worsens agreement with experiment.

Figure 5 shows the anharmonicity correction to the binding free energy of all studied microhydrated monomer, dimer, and (sulfuric acid–ammonia)-pair clusters. The magnitude of the vibrational corrections decrease with increasing system size. With the exception of the multi-region scaling, all scaling approaches reduce the cluster Gibbs free energies to a similar extent, indicating that a single scaling factor provides a reasonable first approximation. In general, applying a single scaling factor systematically lowers reaction and addition free energies, as the Gibbs free energy correction scales approximately linearly with system size. The multi-region scaling correction is significantly larger, which could in principle account for missing contributions to the binding free energies of large clusters. However, when tested for ωB97X-3c, the correction was not significantly more accurate than that obtained with a single scaling factor. The multi-region scaling is thus very sensitive to the data. We also reiterate that this scaling performs worse compared to experimental data.

For comparison, Temelso et al. (2011) studied W_{1-10} clusters at the CCSD(T)/CBS//RI-MP2/aVDZ level of theory and found that the anharmonic correction scales linearly with the number of atoms, amounting to -4.1 kcal mol^{-1} for the 30-atom W_{10} cluster. This suggests that anharmonic effects in some systems or for some methods could be even more significant than those reported here. A more comprehensive treatment of anharmonicity in molecular clusters would therefore be valuable. But, this is theoretically challenging because the multidimensional PES of hydrogen-bonded clusters contains many shallow minima, and coordinate choices strongly influence how anharmonic couplings are represented. Hence, such an analysis lies beyond the scope of the present benchmarking study.

3.3.3 Multi-conformational contribution

In QC, properties are commonly calculated for the lowest free energy conformation, under the assumption that no other low-energy conformers are significantly populated. To assess the importance of multi-conformational entropy, we used the analytical expression of Partanen et al. (2016). According to Partanen et al. (2016), the multi-conformer binding free energy correction ΔG_{mc} is given by:

$$\Delta G_{\mathrm{mc}} = -k_{\mathrm{B}} T \ln \sum_i e^{-\Delta\Delta G_i / k_{\mathrm{B}} T}, \tag{4}$$

where $\Delta\Delta G_i$ denotes the Gibbs binding free energy difference of conformer i relative to the lowest free-energy structure, k_{B} is the Boltzmann constant, and T the temperature. The summation runs over all unique conformers. Although Eq. 4 is written compactly, it carries out full Boltzmann averaging over all conformer energy states: each conformer is weighted by $e^{-\Delta\Delta G_i / k_{\mathrm{B}} T}$, and the logarithmic form simply converts this conformer-weighted partition sum into the corresponding multi-conformer free energy.

Large hydrated clusters can posses multiple significantly populated free energy minima. We, therefore, calculated the multi-conformational free energy contributions using Eq. 4 at ωB97X-D/6-31++G(d,p) for all clusters fulfilling $SA_{0–3}AM_{0–3}W_{0–8}$. See Sec. S7 for more details on the number of minima used. As shown in Fig. 5, the resulting corrections are relatively small, never exceeding -2 kcal mol^{-1}. Moreover, these results do not increase with the number of atoms, indicating that multi-conformational contributions do not become more significant with cluster size. This observation is consistent with the findings of Halonen (2024), who noted that large clusters predominantly occupy a single low-lying minimum.

3.3.4 Umbrella Sampling simulations

In a recent study (Kubečka et al., 2025), we performed umbrella sampling simulations (Torrie and Valleau, 1974) with the PaiNN machine learning potential (Schütt et al., 2019, 2023; Schütt et al., 2021) to calculate cluster Gibbs binding free energies through an approach independent of the statistical thermodynamics traditionally applied in combination with QC results. This approach inherently accounts for anharmonic effects and multi-conformer contributions. Here, we extend this analysis to the $SA_1DMA_1W_n$+W_1 clusters ($n = 0$–4), following the same computational protocol. For additional methodological details, we refer the reader to our earlier work (Kubečka et al., 2025). As shown in Fig 5, the US corrections on top of the QC calculations (red) yields values close to the Halonen limit, suggesting that the traditional QC approach may significantly underestimate cluster binding free energies, and the thermodynamic corrections indeed might be close to the Halonen limit. This methodology was not yet verified but shows that there might be some missing entropic effects in the traditionally applied statistical thermodynamics. In the following sections, we will therefore consider both QC Gibbs binding free energies with QHA corrections, and those additionally corrected according to the Halonen limit ($-k_{\mathrm{B}}T/4$ per vibrational mode). This provides a practical lower and upper bound for the expected binding free energies.

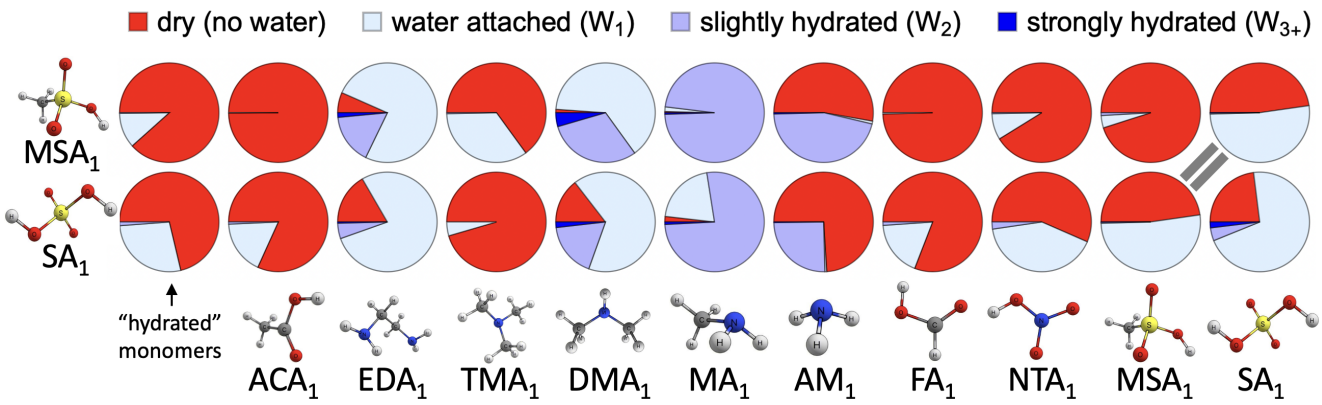

Figure 6. Pie charts of the hydration distribution at 100% relative humidity and 278.15 K for all studied monomers and dimers with up to five water molecules. No hydration is indicated in red, while clusters with one to five water molecules are represented with increasingly darker shades of blue. The results were obtained for the lowest free energy minimum at the DLPNO//DFT level of theory using quasi-harmonic approximation.

3.4 Hydration distribution

Based on the benchmarking of electronic energies, equilibrium geometries, and thermochemical properties in the previous sections, we chose DLPNO$^{\text{NormalPNO}}$–CCSD(T$_0$)/aug-cc-pVTZ//ωB97X-D/6-31++G(d,p), hereafter abreviated as DLPNO//DFT, as the method for calculating hydration distributions. This choice was motivated by the good performance of the DFT method in the equilibrium structure and vibrational analysis benchmarks, while DLPNO excelled in the electronic energy benchmark. Note that this level of theory was already recommended as suitable for molecular clusters in previous studies (Elm et al., 2020; Smith et al., 2021; Trolle et al., 2025). Here, we examine the hydration distributions of various acid–base clusters. First, we consider only the lowest free energy minimum corrected by QHA. Under atmospheric conditions, most monomers and dimers remain predominantly unhydrated (see Sec. S8). Only a few dimers containing sulfuric acid (SA) and methanesulfonic acid (MSA) are more likely to be hydrated than not, as illustrated in Fig. 6. In the figure, unhydrated clusters (i.e., with zero water molecules attached) are shown in red, while clusters with one to five water molecules are represented by increasingly darker shades of blue. The resulting hydration distribution varies with acidity/basicity and structural factors, such as steric hindrance and the availability of hydrogen-bonding sites. Under the same conditions, we also examined the hydration distributions of SA_3AM_3, SA_3DMA_3, and MSA_3MA_3. Figure 7 reveals that the SA_3DMA_3 remains completely dry, which we attribute to the steric effects of the methyl groups. MSA_3MA_3 is the most hydrated among the three, yet it is still less hydrated than MSA_1MA_1, again likely due to the presence of the methyl groups. In contrast, SA_3AM_3 shows increased hydration compared to SA_1AM_1.

The hydration of the SA–AM and SA–DMA systems has previously been studied by Myllys et al. (2021) and Henschel et al. (2014, 2016), and our hydration distributions from QHA-corrected single-structure cluster thermochemistry correspond quite

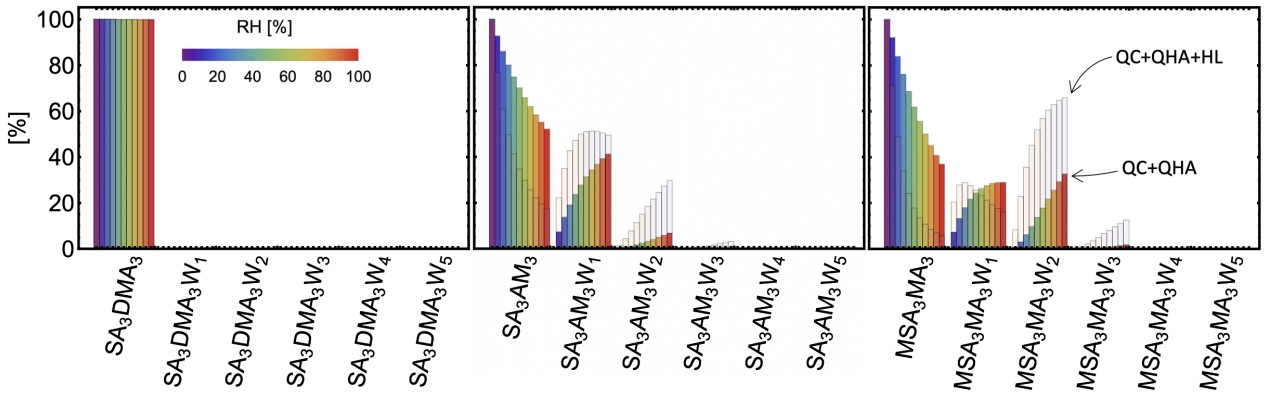

Figure 7. Hydration distributions of the SA_3AM_3, SA_3DMA_3, and MSA_3MA_3 clusters at 278.15 K. The humidity level is indicated by different color. The results were obtained for the lowest free energy minimum at the DLPNO//DFT level of theory using quasi-harmonic approximation. The transparent distribution shows the change due to the free energy correction at the Halonen limit.

well to their results. Chen et al. (2020) further suggested that humidity strongly enhances the NPF of the MSA–MA system, by enabling stabilizing proton transfers. We critically revisit this NPF enhancement in Sec. 3.5. While our results confirm that MSA–MA is more hydrophilic, the effect appears less pronounced than reported by Chen et al. (2020). Additionally, Ge et al. (2020) found that TMA_1 is hydrophobic but, in contrast to our results, reported that DMA_1 is almost always hydrated with one water molecule. We, therefore, examine potential sources of this discrepancy. Rather than temperature dependence (within the atmospheric window), differences in the chosen quantum chemistry methods and thermochemical corrections appear to play a decisive role in shaping the predicted hydration distributions.

As shown in Fig 7, applying the full Halonen-limit correction to the thermochemistry enhances cluster hydration, although the effect remains moderate (cf. the transparent histogram). To examine how the hydration distribution depends on the choice of QC method, both with and without Halonen limit corrections, we calculated the binding free energy of all $(SA_1AM_1)_{1–3}W_{0–5}$ clusters using two different methods: ωB97X-3c and DLPNO//DFT. Additionally, we calculated the hydration distribution for SA–AM clusters up to SA_6AM_6 with up to ten water molecules using the more efficient B97-3c method. We also present the electronically corrected composite methods: LNO//B97-3c, LNO//ωB97X-3c, and DLPNO//ωB97X-3c. Here we omitted DLPNO//B97-3c as it would be slow for the largest clusters. The resulting average numbers of attached water molecules at 278.15 K are presented in Fig. 8. Overall, B97-3c predicts more hydration for SA_1AM_1, SA_2AM_2, and SA_3AM_3 than ωB97X-3c, while ωB97X-3c in turn predicts more hydration than DLPNO//DFT. However, the relative trends are not consistent across cluster sizes. For example, B97-3c predicts reduced hydration when moving from SA_1AM_1 to SA_2AM_2, while DLPNO//DFT predicts an increase in hydration. However, after applying electronic energy corrections to the B97-3c method we observe similar trends. LNO//B97-3c could potentially be a new emerging method for fast calculations. Moreover, similar results are obtained for ωB97X-3c, LNO//ωB97X-3c, and DLPNO//ωB97X-3c, which shows that ωB97X-3c could emerge as a new reasonably accurate method for large clusters without a need for electronic correction. Nevertheless, there are still some dis-

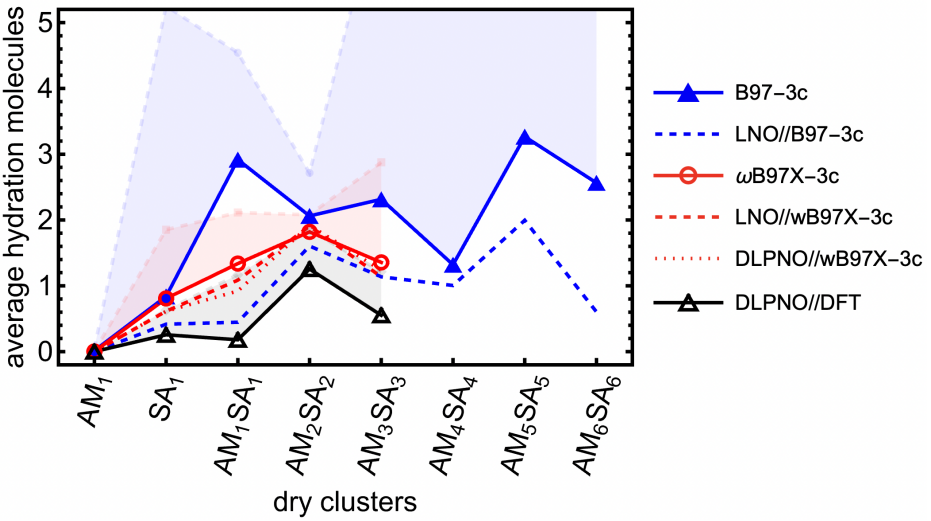

Figure 8. Average hydration of sulfuric acid (SA)–ammonia (AM) clusters calculated at 100% relative humidity and 278.15 K. The colors represents which method was used for geometry optimization and thermochemistry (B97-3c, ωB97X-3c, and DFT), while some of the methods are corrected with single-point calculation at DLPNO$^{\text{NormalPNO}}$–CCSD(T$_0$)/aug-cc-pVTZ and LNO$^{\text{Tight}}$–CCSD(T)/aug$'$-cc-pVTZ level of theory. Bright lines are for calculations with one-structure and quasi-harmonic approximations. The low-opacity line-points corresponds to systematically applied Halonen limit thermodynamics correction to all cluster binding free energies.

crepancies in hydration distributions between methods, and thus relative comparisons should only be made within the same method. Applying the systematic Halonen limit correction further increases the predicted hydration, with the effect being most pronounced for methods that yield stronger binding energies.

When examining the evolution of the hydration distribution with cluster size using the B97-3c method, we find that clusters with even numbers of SA and AM molecules (i.e., SA_2AM_2, SA_4AM_4, SA_6AM_6) are less hydrated than the neighboring clusters with odd numbers of SA and AM molecules. The dry structures with even SA and AM are highly symmetric, resulting in lower energies than the trend in binding free energies with cluster size would suggest. Interestingly, this pattern is not observed with the other two methods. Overall, hydration is rather minor but seems to slowly increase with cluster size because larger clusters can accommodate more water molecules, either by incorporating them into acid–base interaction bridges or by exposing greater surface area for water adsorption.

3.5 Particle Formation Modeling

The initial particle formation rate, J, defined as the rate at which new particles form under given ambient conditions, is the main quantity characterizing the particle formation process (Yazgi and Olenius, 2023). However, the extent to which hydration influences J remains unknown. Figure 9 shows the enhancement of the particle formation rate, $J/J_{\text{RH}=0\%}$, as a function

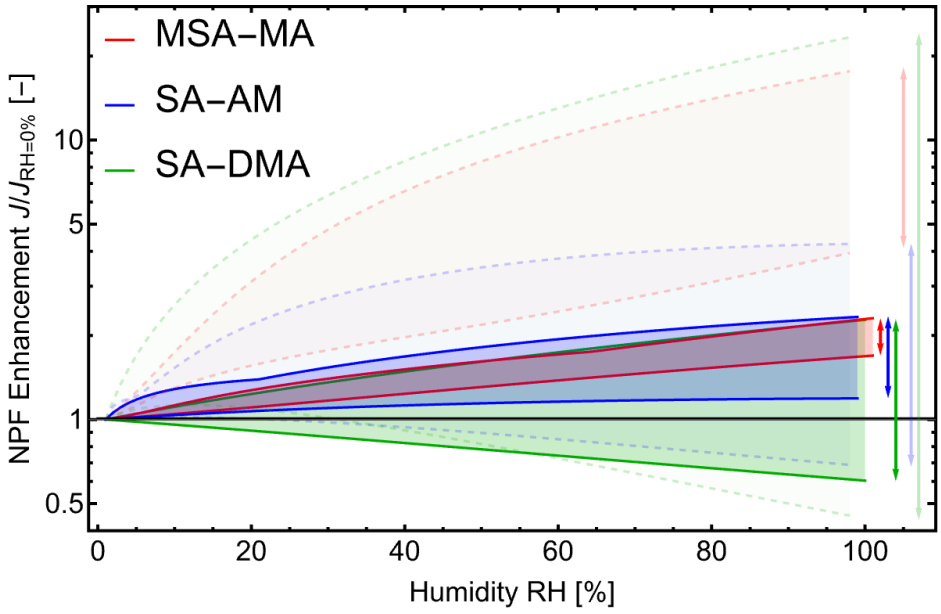

Figure 9. The enhancement in particle formation rate, $J/J_{\mathrm{RH}=0\%}$, as a function of relative humidity (RH), calculated with the Atmospheric Cluster Dynamics Code (ACDC) for clusters up to $\mathrm{acid}_3\mathrm{base}_3\mathrm{water}_5$. Results are shown for the methanesulfonic acid–methylamine (MSA–MA; red), sulfuric acid–ammonia (SA–AM; blue), and sulfuric acid–dimethylamine (SA–DMA; green) systems. Solid lines correspond to DLPNO//DFT quantum-chemical data with QHA, while dashed, semi-transparent lines show results corrected to the Halonen limit. Simulations were performed at 278.15 and 298.15 K, with sulfuric acid concentrations of 10^5 and 10^7 cm^{-3}, methylamine concentrations of 1 and 100 ppt, ammonia concentrations of 10 and 10,000 ppt, and dimethylamine concentrations of 1 and 10 ppt.

of relative humidity (RH). For clarity, the figure highlights only the overall range of enhancement, while detailed results for individual simulations are provided in Sec. S9.

Here we only apply the commonly used and well-performing methodology, $\mathrm{DLPNO}^{\mathrm{NormalPNO}}$–$\mathrm{CCSD(T_0)}$/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory with the quasi-harmonic correction applied, which we use to obtain binding free energies. In addition, we also corrected the binding free energies according to the Halonen limit, providing two limiting cases: without any anharmonic corrections (QC+QHA) and with the largest corrections that can be reasonably expected (QC+QHA+HL). Based on the uncorrected QC+QHA data, the MSA–MA and SA–AM systems exhibit modest, positive enhancements, with J increasing by no more than a factor of two even at high RH. For SA–DMA, hydration can either increase or decrease J depending on the ambient conditions, but the effect remains within a factor of $\sim$2.

When binding free energies are corrected according to the Halonen limit (QC+QHA+HL), we find substantially larger enhancements, with the MSA–MA and SA–DMA systems showing increases of up to a factor of 20. Even so, this corresponds to only about one order of magnitude under the most favorable conditions. Since evaporation rates depend exponentially on binding free energies, an error of just 3 $k_{\mathrm{B}}T$ (≈ 1.8 kcal mol^{-1} at 298 K) can likewise produce a factor of 20 difference in J. As the actual anharmonic corrections are expected to represent only a fraction of the Halonen limit, these results suggest that

hydration does not substantially influence particle formation, particularly when compared to the current level of computational uncertainty across different computational methods.

We note that obtaining quantitatively accurate particle formation rates would require extending the maximum cluster size in the ACDC simulations, as the critical cluster is likely larger than our current cutoff under many of the studied conditions. Consequently, the rates reported here should be interpreted as potential particle formation rates rather than true nucleation rates (Elm, 2021). Nevertheless, we emphasize that our aim is not to determine absolute/potential nucleation rates, but to assess the relative effect of hydration. Because we compare hydrated and dry simulations under otherwise identical conditions, systematic uncertainties largely cancel, making the resulting rate ratios less sensitive to the limited cluster-size cutoff. Thus, while the absolute rates are not quantitative, the relative impact of humidity is more robust.

In comparison with previous studies, our results predict a much smaller enhancement for the MSA–MA system than reported by Chen et al. (2020). At 278.15 K with [MSA] = 10^6 cm^{-3} and [MA] = 10 ppt, they reported enhancements of approximately seven orders of magnitude for $J_{4\times4}$ and three orders of magnitude for $J_{2\times2}$, which can be attributed to a large decrease in binding free energy upon hydration of the clusters. To directly assess this, we recalculated the lowest-energy structures reported in their Supporting Information at the same level of theory (DLPNO$^{\text{NormalPNO}}$–CCSD(T$_0$)/aug-cc-pVTZ//M06-2X/6-31++G(d,p)). We obtained binding energies of -6.25 kcal mol^{-1} for MSA$_1$MA$_1$ and -9.34 kcal mol^{-1} for MSA$_1$MA$_1$W$_1$ at 278.15 K, substantially weaker than the corresponding values of -7.18 kcal mol^{-1} and -13.23 kcal mol^{-1} reported by Chen et al. (2020). While minor differences in DLPNO settings could account for small discrepancies, they cannot explain differences of this magnitude. Moreover, our recalculated values are consistent with those obtained using our DLPNO$^{\text{NormalPNO}}$–CCSD(T$_0$)/aug-cc-pVTZ//ωB97X-D/6-31++G(d,p) method (e.g., -6.53 kcal mol^{-1} and -9.40 kcal mol^{-1}, respectively). Taken together, these results suggest that the hydration effects reported by Chen et al. (2020) for the MSA–MA system are substantially overestimated.

Our results for the SA–DMA system are in good agreement with Henschel et al. (2016), who reported enhancements of at most a factor of two for [SA] = 10^5 cm^{-3} at 263 K. In contrast, Henschel et al. (2016) also found enhancements of up to a factor of 50 for the SA–AM system (263 K, [SA] = 10^5 cm^{-3}, [AM] = 10,000 ppt), which we do not observe under any conditions (even in the Halonen limit used at 263 K).

These comparisons underscore the strong dependence of J on the specific QC calculations employed. To more reliably assess the effects of hydration, calculations should be performed consistently using the same QC methodology across a range of relevant systems.

4 Conclusions

In this work, we systematically benchmarked a broad range of quantum chemistry methods for their ability to describe hydrated atmospheric molecular clusters. The comparison focused on electronic binding energies and equilibrium geometries. Based on this benchmarking, we identified the most accurate methods and used them to evaluate the magnitude of different thermochemical corrections, thereby improving the binding free energies beyond the conventional approach of applying a harmonic

oscillator approximation to a single low-energy structure. Finally, we calculated hydration distributions and particle formation rates using both the uncorrected quantum chemistry binding free energies and the same data adjusted with the maximum correction expected for an ideal system.

The widely used DLPNO$^{\text{NormalPNO}}$–CCSD(T_0)/aug-cc-pVTZ//ωB97X-D/6-31++G(d,p) method performed well in the benchmarks and was therefore employed in the thermochemical analysis as well as in the hydration distribution and particle formation rate calculations, for consistency with previous studies. We also identified ωB97X-3c as an efficient and accurate option for large-scale studies, with or without electronic energy corrections, while DLPNO$^{\text{NormalPNO}}$–CCSD(T_0)/aug-cc-pVTZ and LNO$^{\text{Tight}}$–CCSD(T)/aug$'$-cc-pVTZ provide accurate single-point electronic energy corrections. The low computational cost and memory requirements of the LNO$^{\text{Tight}}$–CCSD(T) methods are particularly noteworthy, and we recommend them for future studies. For very large clusters, B97-3c remains computationally practical, though we advise correcting its electronic energies with LNO$^{\text{Tight}}$–CCSD(T) (Knattrup and Elm, 2025).

We provided a general overview of methods for improving the thermochemical description of molecular clusters, including treatments of vibrational anharmonicity, low-frequency modes, and multi-conformational contributions, each of which can introduce significant corrections to the free energies of hydrated clusters. However, assessing the accuracy of these corrections remains challenging due to the limited availability of experimental reference data. To complement the quantum-chemical approaches, we employed umbrella sampling, which yielded substantially lower binding free energies, approaching the theoretical upper bound on stabilization for ideal systems reported by Halonen (2024). Given the difficulty of determining which methodology provides the most reliable description, we considered a systematic range of cluster free energies and incorporated this uncertainty into the hydration distribution and particle formation rate calculations.

Our results show that the hydration distributions of atmospheric acid–base clusters depend strongly on both cluster composition and the choice of computational method. Because these hydration trends are method-dependent, maintaining consistency in the chosen method is crucial when comparing hydration distributions. Under atmospheric conditions, most monomers and dimers remain largely unhydrated, although dimers containing sulfuric acid or methanesulfonic acid exhibit a greater tendency to bind water. Larger clusters display a pronounced increase in hydration capacity, though the extent depends on factors such as molecular symmetry, available hydrogen-bonding sites, and the presence of hydrophobic alkyl groups. Applying Halonen's upper thermodynamic limit further increases the predicted water content of molecular clusters, though not uniformly across all systems.

When translated into particle formation rates, hydration was found to exert only a modest influence. For sulfuric acid–ammonia, sulfuric acid–dimethylamine, and methanesulfonic acid–methylamine systems, the enhancement of new particle formation (NPF) due to humidity rarely exceeded a factor of $\sim$1–2 under typical atmospheric conditions. In some cases, hydration even suppressed growth by preferentially stabilizing smaller clusters over larger ones. Applying systematic thermochemical corrections (up to the Halonen limit) amplified the effect of hydration, but only up to about one order of magnitude. These results indicate that hydration has a relatively minor thermochemical impact on the earliest steps of NPF. However, this does not imply that humidity is unimportant for later growth, where higher relative humidity may reduce the sticking probability of incoming vapors and where the formation of surface layers at the aerosol–air interface could influence uptake and

stabilization of additional molecules. Future work should extend these benchmarks to larger clusters and later growth stages, where humidity is expected to play a more pronounced role in aerosol evolution.

Comparing to previous studies, our results largely agree with Henschel et al. (2016) regarding the role of hydration in sulfuric acid–ammonia and sulfuric acid–dimethylamine systems. However, the strong humidity dependence reported for methanesulfonic acid–methylamine by Chen et al. (2017) appears to be overstated. While systematic errors in statistical thermodynamics, as discussed by Halonen (2024), do influence absolute predictions of new particle formation rates, their impact on relative humidity-driven enhancements appears to be minor. Nonetheless, the uncertainties associated with statistical thermodynamics following quantum-chemical calculations warrant further attention in future studies.

Data availability. All the calculated structures and thermochemistry are available in the Atmospheric Cluster Database (ACDB) at: https://github.com/elmjonas/ACDB/tree/master/Articles/neefjes25_hydration

Author contributions. Conceptualization: JE, JK; Methodology: IN, HW, JE, JK; Software: JK; Validation: IN, YK, HW, JK; Formal analysis: IN, YK, HW, JK; Investigation: IN, YK, HW, JK; Resources: JE, JK; Data Curation: IN, YK, HW, GBT, JK; Writing - Original Draft: IN, JK; Writing - Review & Editing: IN, YK, HW, GBT, JE, JK; Visualization: IN, JK; Supervision: JE, JK; Project administration: IN, JK; Funding acquisition: JE, JK;

Competing interests. At least one of the (co-)authors is a member of the editorial board of Aerosol Research. The authors have no other competing interests to declare.

Acknowledgements. Funded by the European Union (ERC, ExploreFNP, project 101040353 and MSCA, HYDRO-CLUSTER, project 101105506) and the Danish National Research Foundation (DNRF172) through the Center of Excellence for Chemistry of Clouds. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. The numerical results presented in this work were obtained at the Centre for Scientific Computing, Aarhus https://phys.au.dk/forskning/faciliteter/cscaa/.

References

Alduchov, O. A. and Eskridge, R. E.: Improved Magnus form approximation of saturation vapor pressure, J. Appl. Meteorol. Climatol., 35, 601–609, https://doi.org/10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2, 1996.

Almeida, J., Schobesberger, S., Kürten, A., Ortega, I. K., Kupiainen-Määttä, O., Praplan, A. P., Adamov, A., Amorim, A., Bianchi, F., Breitenlechner, M., David, A., Dommen, J., Donahue, N. M., Downard, A., Dunne, E., Duplissy, J., Ehrhart, S., Flagan, R. C., Franchin, A., Guida, R., Hakala, J., Hansel, A., Heinritzi, M., Henschel, H., Jokinen, T., Junninen, H., Kajos, M., Kangasluoma, J., Keskinen, H., Kupc, A., Kurtén, T., Kvashin, A. N., Laaksonen, A., Lehtipalo, K., Leiminger, M., Leppä, J., Loukonen, V., Makhmutov, V., Mathot, S., McGrath, M. J., Nieminen, T., Olenius, T., Onnela, A., Petäjä, T., Riccobono, F., Riipinen, I., Rissanen, M., Rondo, L., Ruuskanen, T., Santos, F. D., Sarnela, N., Schallhart, S., Schnitzhofer, R., Seinfeld, J. H., Simon, M., Sipilä, M., Stozhkov, Y., Stratmann, F., Tomé, A., Tröstl, J., Tsagkogeorgas, G., Vaattovaara, P., Viisanen, Y., Virtanen, A., Vrtala, A., Wagner, P. E., Weingartner, E., Wex, H., Williamson, C., Wimmer, D., Ye, P., Yli-Juuti, T., Carslaw, K. S., Kulmala, M., Curtius, J., Baltensperger, U., Worsnop, D. R., Vehkamäki, H., and Kirkby, J.: Molecular understanding of sulphuric acid–amine particle nucleation in the atmosphere, Nature, 502, 359–363, https://doi.org/10.1038/nature12663, 2013.

Andersson, M. P.: The shape of water – How cluster formation provides a unifying explanation of water's anomalous properties, J. Mol. Liq., 383, 122 169, https://doi.org/10.1016/j.molliq.2023.122169, 2023.

Andreae, M. O., Talbot, R. W., Andreae, T. W., and Harriss, R. C.: Formic and acetic acid over the central Amazon region, Brazil: 1. dry season, J. Geophys. Res. Atmos., 93, 1616–1624, https://doi.org/10.1029/JD093iD02p01616, 1988.

Bannwarth, C., Ehlert, S., and Grimme, S.: GFN2-xTB–An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions, J. Chem. Theory Comput., 15, 1652–1671, https://doi.org/10.1021/acs.jctc.8b01176, 2019.

Bannwarth, C., Caldeweyher, E., Ehlert, S., Hansen, A., Pracht, P., Seibert, J., Spicher, S., and Grimme, S.: Extended tight-binding quantum chemistry methods, WIREs Comput. Mol. Sci., 11, e1493, https://doi.org/10.1002/wcms.1493, 2021.

Barone, V., Biczysko, M., and Bloino, J.: Fully anharmonic IR and Raman spectra of medium-size molecular systems: accuracy and interpretation, Phys. Chem. Chem. Phys., 16, 1759–1787, https://doi.org/10.1039/C3CP53413H, 2014.

Birmili, W. and Wiedensohler, A.: New particle formation in the continental boundary layer: Meteorological and gas phase parameter influence, Geophys. Res. Lett., 27, 3325–3328, https://doi.org/10.1029/1999GL011221, 2000.

Birmili, W., Berresheim, H., Plass-Dülmer, C., Elste, T., Gilge, S., Wiedensohler, A., and Uhrner, U.: The Hohenpeissenberg aerosol formation experiment (HAFEX): a long-term study including size-resolved aerosol, H_2SO_4, OH, and monoterpenes measurements, Atmos. Chem. Phys., 3, 361–376, https://doi.org/10.5194/acp-3-361-2003, 2003.

Boy, M. and Kulmala, M.: Nucleation events in the continental boundary layer: Influence of physical and meteorological parameters, Atmos. Chem. Phys., 2, 1–16, https://doi.org/10.5194/acp-2-1-2002, 2002.

Brandenburg, J. G., Bannwarth, C., Hansen, A., and Grimme, S.: B97-3c: A revised low-cost variant of the B97-D density functional method, J. Chem. Phys., 148, 064 104, https://doi.org/10.1063/1.5012601, 2018.

Burke, K., Perdew, J. P., and Wang, Y.: Derivation of a generalized gradient approximation: The PW91 density functional, in: Electronic density functional theory: Recent progress and new directions, edited by Dobson, J. F., Vignale, G., and Das, M. P., pp. 81–111, Springer US, Boston, MA, ISBN 978-1-4899-0316-7, https://doi.org/10.1007/978-1-4899-0316-7_7, 1998.

Carlsson, P. T. M., Celik, S., Becker, D., Olenius, T., Elm, J., and Zeuch, T.: Neutral sulfuric acid–water clustering rates: Bridging the gap between molecular simulation and experiment, J. Phys. Chem. Lett., 11, 4239–4244, https://doi.org/10.1021/acs.jpclett.0c01045, 2020.

Chai, J.-D. and Head-Gordon, M.: Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections, Phys. Chem. Chem. Phys., 10, 6615–6620, https://doi.org/10.1039/B810189B, 2008.

Chen, D., Li, D., Wang, C., Luo, Y., Liu, F., and Wang, W.: Atmospheric implications of hydration on the formation of methanesulfonic acid and methylamine clusters: A theoretical study, Chemosphere, 244, 125 538, https://doi.org/10.1016/j.chemosphere.2019.125538, 2020.

Chen, J., Jiang, S., Liu, Y.-R., Huang, T., Wang, C.-Y., Miao, S.-K., Wang, Z.-Q., Zhang, Y., and Huang, W.: Interaction of oxalic acid with dimethylamine and its atmospheric implications, RSC Adv., 7, 6374–6388, https://doi.org/10.1039/C6RA27945G, 2017.

Clark, T., Chandrasekhar, J., Spitznagel, G. W., and Schleyer, P. V. R.: Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li–F, J. Comput. Chem., 4, 294–301, https://doi.org/10.1002/jcc.540040303, 1983.

Coriani, S., Marchesan, D., Gauss, J., Hättig, C., Helgaker, T., and Jørgensen, P.: The accuracy of ab initio molecular geometries for systems containing second-row atoms, J. Chem. Phys, 123, 184 107, https://doi.org/10.1063/1.2104387, 2005.

Dawson, M. L., Varner, M. E., Perraud, V., Ezell, M. J., Gerber, R. B., and Finlayson-Pitts, B. J.: Simplified mechanism for new particle formation from methanesulfonic acid, amines, and water via experiments and ab initio calculations, Proc. Natl. Acad. Sci., 109, 18 719–18 724, https://doi.org/10.1073/pnas.1211878109, 2012.

Del Bene, J. E.: Proton affinities of ammonia, water, and hydrogen fluoride and their anions: A quest for the basis-set limit using the Dunning augmented correlation-consistent basis sets, J. Phys. Chem., 97, 107–110, https://doi.org/10.1021/j100103a020, 1993.

DePalma, J. W., Bzdek, B. R., Doren, D. J., and Johnston, M. V.: Structure and energetics of nanometer size clusters of sulfuric acid with ammonia and dimethylamine, J. Phys. Chem. A, 116, 1030–1040, https://doi.org/10.1021/jp210127w, 2012.

DePalma, J. W., Doren, D. J., and Johnston, M. V.: Formation and growth of molecular clusters containing sulfuric acid, water, ammonia, and dimethylamine, J. Phys. Chem. A, 118, 5464–5473, https://doi.org/10.1021/jp503348b, 2014.

Ditchfield, R., Hehre, W. J., and Pople, J. A.: Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules, J. Chem. Phys., 54, 724–728, https://doi.org/10.1063/1.1674902, 1971.

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

Dunning, T. H.: Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys., 90, 1007–1023, https://doi.org/10.1063/1.456153, 1989.

Duplissy, J., Merikanto, J., Franchin, A., Tsagkogeorgas, G., Kangasluoma, J., Wimmer, D., Vuollekoski, H., Schobesberger, S., Lehtipalo, K., Flagan, R. C., Brus, D., Donahue, N. M., Vehkamäki, H., Almeida, J., Amorim, A., Barmet, P., Bianchi, F., Breitenlechner, M., Dunne, E. M., Guida, R., Henschel, H., Junninen, H., Kirkby, J., Kürten, A., Kupc, A., Määttänen, A., Makhmutov, V., Mathot, S., Nieminen, T., Onnela, A., Praplan, A. P., Riccobono, F., Rondo, L., Steiner, G., Tome, A., Walther, H., Baltensperger, U., Carslaw, K. S., Dommen, J., Hansel, A., Petäjä, T., Sipilä, M., Stratmann, F., Vrtala, A., Wagner, P. E., Worsnop, D. R., Curtius, J., and Kulmala, M.: Effect of ions on sulfuric acid-water binary particle formation: 2. Experimental data and comparison with QC-normalized classical nucleation theory, J. Geophys. Res. Atmos., 121, 1752–1775, https://doi.org/10.1002/2015JD023539, 2016.

Elm, J.: An atmospheric cluster database consisting of sulfuric acid, bases, organics, and water, ACS Omega, 4, 10 965–10 974, https://doi.org/10.1021/acsomega.9b00860, 2019.

Elm, J.: Clusteromics I: Principles, Protocols, and Applications to Sulfuric Acid–Base Cluster Formation, ACS Omega, 6, 7804–7814, https://doi.org/10.1021/acsomega.1c00306, 2021.

Elm, J. and Kubečka, J.: Atmospheric Cluster Database (ACDB), https://doi.org/10.1021/acsomega.9b00860, 2024.

Elm, J. and Mikkelsen, K. V.: Computational approaches for efficiently modelling of small atmospheric clusters, Chem. Phys. Lett., 615, 26–29, https://doi.org/10.1016/j.cplett.2014.09.060, 2014.

Elm, J., Kubečka, J., Besel, V., Jääskeläinen, M. J., Halonen, R., Kurtén, T., and Vehkamäki, H.: Modeling the formation and growth of atmospheric molecular clusters: A review, J. Aerosol Sci., 149, 105 621, https://doi.org/10.1016/j.jaerosci.2020.105621, 2020.

Engsvang, M. and Elm, J.: Modeling the binding free energy of large atmospheric sulfuric acid–ammonia clusters, ACS Omega, 7, 8077–8083, https://doi.org/10.1021/acsomega.1c07303, 2022.

Falcon-Rodriguez, C. I., Osornio-Vargas, A., Sada-Ovalle, I., and Segura-Medina, P.: Aeroparticles, composition, and lung diseases, Front. Immunol., 7, 1–9, https://doi.org/10.3389/fimmu.2016.00003, 2016.

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

Fernández, L., Marigliano, A. G., and Varetti, E.: The vibrational properties of formic acid as monomer and dimer: A DFT study, Vib. Spectrosc., 37, 179–187, https://doi.org/j.vibspec.2004.09.001, 2005.

Francl, M. M., Pietro, W. J., Hehre, W. J., Binkley, J. S., Gordon, M. S., DeFrees, D. J., and Pople, J. A.: Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements, J. Chem. Phys., 77, 3654–3665, https://doi.org/10.1063/1.444267, 1982.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B., and Fox, D. J.: Gaussian 16, Revision B.01, available at https://gaussian.com/g16/, 2016.

Galloway, J. N., Likens, G. E., Keene, W. C., and Miller, J. M.: The composition of precipitation in remote areas of the world, J. Geophys. Res. Oceans, 87, 8771–8786, https://doi.org/10.1029/JC087iC11p08771, 1982.

Gao, Y., Fang, H., and Ni, K.: Water clusters and density fluctuations in liquid water based on extended hierarchical clustering methods, Sci. Rep., 12, 8036, https://doi.org/10.1038/s41598-022-11947-6, 2022.

García-Argote, W., Ruiz, L., Inostroza, D., Cardenas, C., Yañez, O., and Tiznado, W.: Introducing KICK-MEP: Exploring potential energy surfaces in systems with significant non-covalent interactions, J. Mol. Model., 30, 369, https://doi.org/10.1007/s00894-024-06155-0, 2024.

Ge, P., Luo, G., Huang, W., Xie, H., Chen, J., and Luo, Y.: Theoretical study of the hydration effects on alkylamine and alkanolamine clusters and the atmospheric implication, Chemosphere, 243, 125 323, https://doi.org/https://doi.org/10.1016/j.chemosphere.2019.125323, 2020.

Gong, X., Zhu, L., and Zhang, R.: Technical note: A theoretical study on the mechanism of citric acid-driven multi-component nucleation of sulfuric acid-base-water clusters, EGUsphere, 2024, 1–18, https://doi.org/10.5194/egusphere-2023-3113, 2024.

Gordon, M. S., Binkley, J. S., Pople, J. A., Pietro, W. J., and Hehre, W. J.: Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements, J. Am. Chem. Soc., 104, 2797–2803, https://doi.org/10.1021/ja00374a017, 1982.

Grimme, S.: Supramolecular binding thermodynamics by dispersion-corrected density functional theory, Chem. Eur. J., 18, 9955–9964, https://doi.org/10.1002/chem.201200497, 2012.

Grimme, S., Bannwarth, C., and 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., 13, 1989–2009, https://doi.org/10.1021/acs.jctc.7b00118, 2017.

Grimme, S., Hansen, A., Ehlert, S., and Mewes, J.-M.: r^2SCAN-3c: A "Swiss army knife" composite electronic-structure method, J. Chem. Phys., 154, 064 103, https://doi.org/10.1063/5.0040021, 2021.

Halonen, R.: Assessment of anharmonicities in clusters: Developing and validating a minimum-information partition function, J. Chem. Theory Comput., 20, 4099–4114, https://doi.org/10.1021/acs.jctc.4c00121, 2024.

Hamed, A., Korhonen, H., Sihto, S.-L., Joutsensaari, J., Järvinen, H., Petäjä, T., Arnold, F., Nieminen, T., Kulmala, M., Smith, J. N., Lehtinen, K. E. J., and Laaksonen, A.: The role of relative humidity in continental new particle formation, J. Geophys. Res. Atmos., 116, D03 202, https://doi.org/10.1029/2010JD014186, 2011.

Hariharan, P. C. and Pople, J. A.: The influence of polarization functions on molecular orbital hydrogenation energies, Theor. Chim. Acta, 28, 213–222, https://doi.org/10.1007/bf00533485, 1973.

Haywood, J. and Boucher, O.: Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A Review, Rev. Geophys., 38, 513–543, https://doi.org/10.1029/1999RG000078, 2020.

Hehre, W. J., Ditchfield, R., and Pople, J. A.: Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules, J. Chem. Phys., 56, 2257–2261, https://doi.org/10.1063/1.1677527, 1972.

Henschel, H., Navarro, J. C., Yli-Juuti, T., Kupiainen-Määttä, O., Olenius, T., Ortega, I. K., Clegg, S. L., Kurtén, T., Riipinen, I., and Vehkamäki, H.: Hydration of atmospherically relevant molecular clusters: Computational chemistry and classical thermodynamics, J. Phys. Chem. A, 118, 2599–2611, https://doi.org/10.1021/jp500712y, 2014.

Henschel, H., Kurtén, T., and Vehkamäki, H.: Computational study on the effect of hydration on new particle formation in the sulfuric acid/ammonia and sulfuric acid/dimethylamine systems, J. Phys. Chem. A, 120, 1886–1896, https://doi.org/10.1021/acs.jpca.5b11366, 2016.

Herzberg, G.: Electronic Spectra and Electronic Structure of Polyatomic Molecules, Molecular spectra and molecular structure, Van Nostrand, ISBN 9780442033873, https://books.google.dk/books?id=it8NAQAAIAAJ, 1966.

Hintze, P. E., Kjaergaard, H. G., Vaida, V., and Burkholder, J. B.: Vibrational and electronic spectroscopy of sulfuric acid vapor, J. Phys. Chem. A., 107, 1112–1118, https://doi.org/10.1021/jp0263626, 2003.

Hu, Y.-C., Zhang, X.-H., Li, Q.-S., Zhang, Y.-H., and Li, Z.-S.: Effect of water on the structure and stability of hydrogen-bonded oxalic acid dimer, Chem. Phys. Chem., 18, 3375–3383, https://doi.org/10.1002/cphc.201700950, 2017.

Huber, K. P. and Herzberg, G.: Constants of diatomic molecules, pp. 8–689, Springer US, Boston, MA, ISBN 978-1-4757-0961-2, https://doi.org/10.1007/978-1-4757-0961-2_2, 1979.

Ianni, J. and Bandy, A.: A theoretical study of the hydrates of $(H_2SO_4)_2$ and its implications for the formation of new atmospheric particles, J. Mol. Struct. THEOCHEM, 497, 19–37, https://doi.org/10.1016/S0166-1280(99)00182-7, 2000.

Intergovernmental Panel on Climate Change (IPCC): Climate Change 2021 – The physical science basis: Working group I Contribution to the sixth assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2023.

Jacobsen, R. L., Johnson, R. D. I., Irikura, K. K., and Kacker, R. N.: Anharmonic vibrational frequency calculations are not worthwhile for small basis sets, J. Chem. Theory Comput., 9, 951–954, https://doi.org/10.1021/ct300293a, 2013.

Jen, C. N., Bachman, R., Zhao, J., McMurry, P. H., and Hanson, D. R.: Diamine-sulfuric acid reactions are a potent source of new particle formation, Geophys. Res. Lett., 43, 867–873, https://doi.org/10.1002/2015GL066958, 2016.

Jensen, A. B. and Elm, J.: Massive assessment of the geometries of atmospheric molecular clusters, J. Chem. Theory Comput., 20, 8549–8558, https://doi.org/10.1021/acs.jctc.4c01046, 2024.

Jensen, A. B., Kubečka, J., Schmitz, G., Christiansen, O., and Elm, J.: Massive assessment of the binding energies of atmospheric molecular clusters, J. Chem. Theory Comput., 18, 7373–7383, https://doi.org/10.1021/acs.jctc.2c00825, 2022.

Joel M Bowman, T. C. and Meyer, H.-D.: Variational quantum approaches for computing vibrational energies of polyatomic molecules, Mol. Phys., 106, 2145–2182, https://doi.org/10.1080/00268970802258609, 2008.

Johnson, R.: NIST 101. Computational Chemistry Comparison and Benchmark Database, 1999.

Jokinen, T., Sipilä, M., Junninen, H., Ehn, M., Lönn, G., Hakala, J., Petäjä, T., Mauldin III, R. L., Kulmala, M., and Worsnop, D. R.: Atmospheric sulphuric acid and neutral cluster measurements using CI-APi-TOF, Atmos. Chem. Phys., 12, 4117–4125, https://doi.org/10.5194/acp-12-4117-2012, 2012.

Keene, W. C. and Galloway, J. N.: Organic acidity in precipitation of North America, Atmos. Environ. (1967), 18, 2491–2497, https://doi.org/10.1016/0004-6981(84)90020-9, 1984.

Keene, W. C., Galloway, J. N., and Holden Jr., J. D.: Measurement of weak organic acidity in precipitation from remote areas of the world, J. Geophys. Res. Oceans, 88, 5122–5130, https://doi.org/10.1029/JC088iC09p05122, 1983.

Kendall, R. A., Dunning, T. H., and Harrison, R. J.: Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, J. Chem. Phys., 96, 6796–6806, https://doi.org/10.1063/1.462569, 1992.

Kildgaard, J. V., Mikkelsen, K. V., Bilde, M., and Elm, J.: Hydration of atmospheric molecular clusters: A new method for systematic configurational sampling, J. Phys. Chem. A, 122, 5026–5036, https://doi.org/10.1021/acs.jpca.8b02758, 2018a.

Kildgaard, J. V., Mikkelsen, K. V., Bilde, M., and Elm, J.: Hydration of atmospheric molecular clusters II: Organic acid–water clusters, J. Phys. Chem. A, 122, 8549–8556, https://doi.org/10.1021/acs.jpca.8b07713, 2018b.

Kirkby, J., Curtius, J., Almeida, J., Dunne, E., Duplissy, J., Ehrhart, S., Franchin, A., Gagné, S., Ickes, L., Kürten, A., and et al.: Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation, Nature, 476, 429–433, https://doi.org/10.1038/nature10343, 2011.

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

Knattrup, Y. and Elm, J.: Extrapolating local coupled cluster calculations toward CCSD(T)/CBS binding energies of atmospheric molecular clusters, ChemRxiv, https://doi.org/10.26434/chemrxiv-2025-704l3-v2, 2025.

Knattrup, Y., Kubečka, J., Wu, H., Jensen, F., and Elm, J.: Reparameterization of GFN1-xTB for atmospheric molecular clusters: Applications to multi-acid–multi-base systems, RSC Adv., 14, 20 048–20 055, https://doi.org/10.1039/D4RA03021D, 2024.

Kodrycka, M. and Patkowski, K.: Platinum, gold, and silver standards of intermolecular interaction energy calculations, J. Chem. Phys., 151, 070 901, https://doi.org/10.1063/1.5116151, 2019.

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

Krishnan, R., Binkley, J. S., Seeger, R., and Pople, J. A.: Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions, J. Chem. Phys., 72, 650–654, https://doi.org/10.1063/1.438955, 1980.

Kruse, H., Szabla, R., and Šponer, J.: Surprisingly broad applicability of the cc-pVnZ-F12 basis set for ground and excited states, J. Chem. Phys., 152, 214 104, https://doi.org/10.1063/5.0006871, 2020.

Kubečka, J., Besel, V., Neefjes, I., Knattrup, Y., Kurtén, T., Vehkamäki, H., and Elm, J.: Computational tools for handling molecular clusters: Configurational sampling, storage, analysis, and machine learning, ACS Omega, 8, 45 115–45 128, https://doi.org/10.1021/acsomega.3c07412, 2023.

Kubečka, J., Knattrup, Y., Trolle, G. B., Reischl, B., Lykke-Møller, A. S., Elm, J., and Neefjes, I.: Thermodynamics of molecular binding and clustering in the atmosphere revealed through conventional and ML-enhanced umbrella sampling, ACS Omega, 10, 39 148–39 161, https://doi.org/10.1021/acsomega.5c05634, 2025.

Kulmala, M., Kontkanen, J., Junninen, H., Lehtipalo, K., Manninen, H. E., Nieminen, T., Petäjä, T., Sipilä, M., Schobesberger, S., Rantala, P., Franchin, A., Jokinen, T., Järvinen, E., Äijälä, M., Kangasluoma, J., Hakala, J., Aalto, P. P., Paasonen, P., Mikkilä, J., Vanhanen, J., Aalto, J., Hakola, H., Makkonen, U., Ruuskanen, T., Mauldin, R. L., Duplissy, J., Vehkamäki, H., Bäck, J., Kortelainen, A., Riipinen, I., Kurtén, T., Johnston, M. V., Smith, J. N., Ehn, M., Mentel, T. F., Lehtinen, K. E. J., Laaksonen, A., Kerminen, V.-M., and Worsnop, D. R.: Direct Observations of Atmospheric Aerosol Nucleation, Science, 339, 943–946, https://doi.org/10.1126/science.1227385, 2013.

Kurtén, T., Loukonen, V., Vehkamäki, H., and Kulmala, M.: Amines are likely to enhance neutral and ion-induced sulfuric acid-water nucleation in the atmosphere more effectively than ammonia, Atmos. Chem. Phys., 8, 4095–4103, https://doi.org/10.5194/acp-8-4095-2008, 2008.

Kurtén, T., Torpo, L., Ding, C.-G., Vehkamäki, H., Sundberg, M. R., Laasonen, K., and Kulmala, M.: A density functional study on water-sulfuric acid-ammonia clusters and implications for atmospheric cluster formation, J. Geophys. Res. D: Atmos., 112, D04 210, https://doi.org/10.1029/2006JD007391, 2007.

Kállay, M., Nagy, P. R., Mester, D., Rolik, Z., Samu, G., Csontos, J., Csóka, J., Szabó, P. B., Gyevi-Nagy, L., Hégely, B., Ladjánszki, I., Szegedy, L., Ladóczki, B., Petrov, K., Farkas, M., Mezei, P. D., and Ganyecz, Á.: The MRCC program system: Accurate quantum chemistry from water to proteins, J. Chem. Phys., 152, 074 107, https://doi.org/10.1063/1.5142048, 2020.

Kállay, M., Nagy, P. R., Mester, D., Gyevi-Nagy, L., Csóka, J., Szabó, P. B., Rolik, Z., Samu, G., Hégely, B., Ladóczki, B., Petrov, K., Csontos, J., Ganyecz, Á., Ladjánszki, I., Szegedy, L., Farkas, M., Mezei, P. D., Horváth, R. A., and Lőrincz, B. D.: MRCC: A quantum chemical program suite, http://www.mrcc.hu, 2025.

Laaksonen, A., Kulmala, M., O'Dowd, C. D., Joutsensaari, J., Vaattovaara, P., Mikkonen, S., Lehtinen, K. E. J., Sogacheva, L., Dal Maso, M., Aalto, P., Petäjä, T., Sogachev, A., Yoon, Y. J., Lihavainen, H., Nilsson, D., Facchini, M. C., Cavalli, F., Fuzzi, S., Hoffmann, T., Arnold, F., Hanke, M., Sellegri, K., Umann, B., Junkermann, W., Coe, H., Allan, J. D., Alfarra, M. R., Worsnop, D. R., Riekkola, M.-L., Hyötyläinen, T., and Viisanen, Y.: The role of VOC oxidation products in continental new particle formation, Atmos. Chem. Phys., 8, 2657–2665, https://doi.org/10.5194/acp-8-2657-2008, 2008.

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

Li, J., Carlson, B. E., Yung, Y. L., Lv, D., Hansen, J., Penner, J. E., Liao, H., Ramaswamy, V., Kahn, R. A., Zhang, P., and et al.: Scattering and absorbing aerosols in the climate system, Nat. Rev. Earth Environ., 3, 363–379, https://doi.org/10.1038/s43017-022-00296-7, 2022.

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

Lin, C., Gilbert, A., and Gill, P.: Calculating molecular vibrational spectra beyond the harmonic approximation, Theor. Chem. Acc., 120, 23–35, https://doi.org/10.1007/s00214-007-0292-8, 2008.

Liu, L., Li, H., Zhang, H., Zhong, J., Bai, Y., Ge, M., Li, Z., Chen, Y., and Zhang, X.: The role of nitric acid in atmospheric new particle formation, Phys. Chem. Chem. Phys., 20, 17 406–17 414, https://doi.org/10.1039/C8CP02719F, 2018.

Lohmann, U. and Feichter, J.: Global indirect aerosol effects: A review, Atmos. Phys. Chem., 5, 715–737, https://doi.org/10.5194/acp-5-715-2005, 2005.

Maroń, M. K., Shultz, M. J., and 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., 473, 268–273, https://doi.org/10.1016/j.cplett.2009.03.071, 2009.

Maso, M. D., Hyvärinen, A., Komppula, M., Tunved, P., Kerminen, V., Lihavainen, H., Öviisanen, Y., Hansson, H.-C., and 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., 60, 495–508, https://doi.org/10.1111/j.1600-0889.2008.00366.x, 2008.

McCurdy, P. R., Hess, W. P., and Xantheas, S. S.: Nitric acid–water complexes: Theoretical calculations and comparison to experiment, J. Phys. Chem. A., 106, 7628–7635, https://doi.org/10.1021/jp020257e, 2002.

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

McLean, A. D. and Chandler, G. S.: Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z=11–18, J. Chem. Phys., 72, 5639–5648, https://doi.org/10.1063/1.438980, 1980.

McMurry, P. H.: The history of condensation nucleus counters, Aerosol Sci. Technol., 33, 297–322, https://doi.org/10.1080/02786820050121512, 2000.

Mei, M., Song, H., Chen, L., Hu, B., Bai, R., Xu, D., Liu, Y., Zhao, Y., and Chen, C.: Early-life exposure to three size-fractionated ultrafine and fine atmospheric particulates in Beijing exacerbates asthma development in mature mice, Part. Fibre Toxicol., 15, 13, https://doi.org/10.1186/s12989-018-0249-1, 2018.

Merikanto, J., Duplissy, J., Määttänen, A., Henschel, H., Donahue, N. M., Brus, D., Schobesberger, S., Kulmala, M., and Vehkamäki, H.: Effect of ions on sulfuric acid-water binary particle formation: 1. Theory for kinetic- and nucleation-type particle formation and atmospheric implications, J. Geophys. Res. Atmos., 121, 1736–1751, https://doi.org/10.1002/2015JD023538, 2016.

Miao, S.-K., Jiang, S., Chen, J., Ma, Y., Zhu, Y.-P., Wen, Y., Zhang, M.-M., and Huang, W.: Hydration of a sulfuric acid–oxalic acid complex: Acid dissociation and its atmospheric implication, RSC Adv., 5, 48 638–48 646, https://doi.org/10.1039/C5RA06116D, 2015.

Millet, D. B., Baasandorj, M., Farmer, D. K., Thornton, J. A., Baumann, K., Brophy, P., Chaliyakunnel, S., de Gouw, J. A., Graus, M., Hu, L., Koss, A., Lee, B. H., Lopez-Hilfiker, F. D., Neuman, J. A., Paulot, F., Peischl, J., Pollack, I. B., Ryerson, T. B., Warneke, C., Williams, B. J., and Xu, J.: A large and ubiquitous source of atmospheric formic acid, Atmos. Chem. Phys., 15, 6283–6304, https://doi.org/10.5194/acp-15-6283-2015, 2015.

Myllys, N.: The role of hydration in atmospheric salt particle formation, Phys. Chem. Chem. Phys., 25, 7394–7400, https://doi.org/10.1039/D3CP00049D, 2023.

Myllys, N., Elm, J., and Kurtén, T.: Density functional theory basis set convergence of sulfuric acid-containing molecular clusters, Comput. Theor. Chem., 1098, 1–12, https://doi.org/10.1016/j.comptc.2016.10.015, 2016.

Myllys, N., Kubečka, J., Besel, V., Alfaouri, D., Olenius, T., Smith, J. N., and Passananti, M.: Role of base strength, cluster structure and charge in sulfuric-acid-driven particle formation, Atmos. Chem. Phys., 19, 9753–9768, https://doi.org/10.5194/acp-19-9753-2019, 2019.

Myllys, N., Myers, D., Chee, S., and Smith, J. N.: Molecular properties affecting the hydration of acid–base clusters, Phys. Chem. Chem. Phys., 23, 13 106–13 114, https://doi.org/10.1039/D1CP01704G, 2021.

Müller, M., Hansen, A., and Grimme, S.: ωB97X-3c: A composite range-separated hybrid DFT method with a molecule-optimized polarized valence double-ζ basis set, J. Chem. Phys., 158, 014 103, https://doi.org/10.1063/5.0133026, 2023.

Nadykto, A. B. and Yu, F.: Strong hydrogen bonding between atmospheric nucleation precursors and common organics, Chem. Phys. Lett., 435, 14–18, https://doi.org/10.1016/j.cplett.2006.12.050, 2007.

Nagy, P. R. and Kállay, M.: Approaching the basis set limit of CCSD(T) energies for large molecules with local natural orbital coupled-cluster methods, J. Chem. Theory Comput., 15, 5275–5298, https://doi.org/10.1021/acs.jctc.9b00511, 2019.

Nagy, P. R., Samu, G., and Kállay, M.: Optimization of the linear-scaling local natural orbital CCSD(T) method: Improved algorithm and benchmark applications, J. Chem. Theory Comput., 14, 4193–4215, https://doi.org/10.1021/acs.jctc.8b00442, 2018.

Nandi, A., Qu, C., Houston, P. L., Conte, R., Yu, Q., and Bowman, J. M.: A CCSD(T)-based 4-body potential for water, J. Phys. Chem. Lett., 12, 10 318–10 324, https://doi.org/10.1021/acs.jpclett.1c03152, 2021.

Neese, F.: The ORCA program system, WIREs Comput. Mol. Sci., 2, 73–78, https://doi.org/10.1002/wcms.81, 2012.

Neese, F.: Software update: The ORCA program system, version 5.0, WIREs Comput. Mol. Sci., 12, e1606, https://doi.org/10.1002/wcms.1606, 2022.

Nguyen, Q. C., Ong, Y. S., Soh, H., and Kuo, J.-L.: Multiscale approach to explore the potential energy surface of water clusters $(H_2O)_n$ $n \leq 8$, J. Phys. Chem. A, 112, 6257–6261, https://doi.org/10.1021/jp802118j, 2008.

Odbadrakh, T. T., Gale, A. G., Ball, B. T., Temelso, B., and Shields, G. C.: Computation of atmospheric concentrations of molecular clusters from ab initio thermochemistry, JoVE, p. e60964, https://doi.org/doi:10.3791/60964, 2020.

Olenius, T.: ACDC: Atmospheric Cluster Dynamics Code, https://github.com/tolenius/ACDC, 2018.

Olenius, T., Kupiainen-Määttä, O., Ortega, I. K., Kurtén, T., and Vehkamäki, H.: Free energy barrier in the growth of sulfuric acid–ammonia and sulfuric acid–dimethylamine clusters, J. Chem. Phys., 139, 084 312, https://doi.org/10.1063/1.4819024, 2013.

Otto, K. E., Xue, Z., Zielke, P., and Suhm, M. A.: The Raman spectrum of isolated water clusters, Phys. Chem. Chem. Phys., 16, 9849–9858, https://doi.org/10.1039/C3CP54272F, 2014.

Partanen, L., Vehkamäki, H., Hansen, K., Elm, J., Henschel, H., Kurtén, T., Halonen, R., and Zapadinsky, E.: Effect of conformers on free energies of atmospheric complexes, J. Phys. Chem. A., 120, 8613–8624, https://doi.org/10.1021/acs.jpca.6b04452, 2016.

Passananti, M., Zapadinsky, E., Zanca, T., Kangasluoma, J., Myllys, N., Rissanen, M. P., Kurtén, T., Ehn, M., Attoui, M., and Vehkamäki, H.: How well can we predict cluster fragmentation inside a mass spectrometer?, Chem. Commun., 55, 5946–5949, https://doi.org/10.1039/C9CC02896J, 2019.

Pavošević, F., Peng, C., Pinski, P., Riplinger, C., Neese, F., and Valeev, E. F.: SparseMaps—A systematic infrastructure for reduced scaling electronic structure methods. V. Linear scaling explicitly correlated coupled-cluster method with pair natural orbitals, J. Chem. Phys, 146, 174 108, https://doi.org/10.1063/1.4979993, 2017.

Peterson, K. A., Adler, T. B., and Werner, H.-J.: Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B–Ne, and Al–Ar, J. Chem. Phys., 128, 084 102, https://doi.org/10.1063/1.2831537, 2008.

Pitman, S. J., Evans, A. K., Ireland, R. T., Lempriere, F., and McKemmish, L. K.: Benchmarking basis sets for density functional theory thermochemistry calculations: Why unpolarized basis sets and the polarized 6-311G family should be avoided, J. Phys. Chem. A, 127, 10 295–10 306, https://doi.org/10.1021/acs.jpca.3c05573, 2023.

Ramabhadran, R. O. and Raghavachari, K.: Extrapolation to the gold-standard in quantum chemistry: Computationally efficient and accurate CCSD(T) energies for large molecules using an automated thermochemical hierarchy, J. Chem. Theory Comput., 9, 3986–3994, https://doi.org/10.1021/ct400465q, 2013.

Rasmussen, F. R., Kubečka, J., Besel, V., Vehkamäki, H., Mikkelsen, K. V., Bilde, M., and Elm, J.: Hydration of atmospheric molecular clusters III: Procedure for efficient free energy surface exploration of large hydrated clusters, J. Phys. Chem. A., 124, 5253–5261, https://doi.org/10.1021/acs.jpca.0c02932, 2020.

Rasmussen, F. R., Kubečka, J., and Elm, J.: Contribution of methanesulfonic acid to the formation of molecular clusters in the marine atmosphere, J. Phys. Chem. A., 126, 7127–7136, https://doi.org/10.1021/acs.jpca.2c04468, 2022.

Riplinger, C. and Neese, F.: An efficient and near linear scaling pair natural orbital based local coupled cluster method, J. Chem. Phys., 138, 034 106, https://doi.org/10.1063/1.4773581, 2013.

Rognoni, A., Conte, R., and Ceotto, M.: How many water molecules are needed to solvate one?, Chem. Sci., 12, 2060–2064, https://doi.org/10.1039/D0SC05785A, 2021.

Rolik, Z., Szegedy, L., Ladjánszki, I., Ladóczki, B., and Kállay, M.: An efficient linear-scaling CCSD(T) method based on local natural orbitals, J. Chem. Phys., 139, 094 105, https://doi.org/10.1063/1.4819401, 2013.

Rozenberg, M., Loewenschuss, A., and Nielsen, C. J.: H-bonded clusters in the trimethylamine/water system: A matrix isolation and computational study, J. Phys. Chem. A., 116, 4089–4096, https://doi.org/10.1021/jp3020035, 2012.

Schmitz, G. and Elm, J.: Assessment of the DLPNO binding energies of strongly noncovalent bonded atmospheric molecular clusters, ACS Omega, 5, 7601–7612, https://doi.org/10.1021/acsomega.0c00436, 2020.

Schütt, K., Unke, O., and Gastegger, M.: Equivariant message passing for the prediction of tensorial properties and molecular spectra, Int. Conf. Mach. Learn., pp. 9377–9388, 2021.

Schütt, K. T., Kessel, P., Gastegger, M., Nicoli, K. A., Tkatchenko, A., and Müller, K.: SchNetPack: A deep learning toolbox for atomistic systems, J. Chem. Theory Comput., 15, 448–455, https://doi.org/10.1021/acs.jctc.8b00908, 2019.

Schütt, K. T., Hessmann, S. S. P., Gebauer, N. W. A., Lederer, J., and Gastegger, M.: SchNetPack 2.0: A neural network toolbox for atomistic machine learning, J. Chem. Phys., 158, 144 801, https://doi.org/10.1063/5.0138367, 2023.

Shimanouchi, T., Shimanouchi, T., and of Standards, U. S. N. B.: Tables of Molecular Vibrational Frequencies. Consolidated Volume I, National Standard Reference Data Series, U.S. Government, 1972.

Sipilä, M., Berndt, T., Petäjä, T., Brus, D., Vanhanen, J., Stratmann, F., Patokoski, J., III, R. L. M., Hyvärinen, A.-P., Lihavainen, H., and Kulmala, M.: The role of sulfuric acid in atmospheric nucleation, Science, 327, 1243–1246, https://doi.org/10.1126/science.1180315, 2010.

Smith, J. N., Draper, D. C., Chee, S., Dam, M., Glicker, H., Myers, D., Thomas, A. E., Lawler, M. J., and Myllys, N.: Atmospheric clusters to nanoparticles: Recent progress and challenges in closing the gap in chemical composition, J. Aerosol Sci., 153, 105 733, https://doi.org/https://doi.org/10.1016/j.jaerosci.2020.105733, 2021.

Soulard, P. and Tremblay, B.: Vibrational study of methylamine dimer and hydrated methylamine complexes in solid neon supported by ab initio calculations, J. Mol. Struct., 1236, 130 308, https://doi.org/10.1016/j.molstruc.2021.130308, 2021.

Spitznagel, G. W., Clark, T., Schleyer, P. v. R., and Hehre, W. J.: An evaluation of the performance of diffuse function-augmented basis sets for second row elements, Na–Cl, J. Comput. Chem., 8, 1109–1116, https://doi.org/10.1002/jcc.540080807, 1987.

Stewart, J. J.: Optimization of parameters for semiempirical methods VI: More modifications to the NDDO approximations and re-optimization of parameters, J. Mol. Model., 19, 1–32, https://doi.org/10.1007/s00894-012-1667-x, 2012.

Telfah, A., Charifi, Z., Latelli, N., Qattan, I. A., Baaziz, H., Al-Bataineh, Q. M., Alsaad, A., and 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., 14, 11 252, https://doi.org/10.1038/s41598-024-61364-0, 2024.

Temelso, B., Archer, K. A., and Shields, G. C.: Benchmark structures and binding energies of small water clusters with anharmonicity corrections, J. Phys. Chem. A, 115, 12 034–12 046, https://doi.org/10.1021/jp2069489, 2011.

Temelso, B., Morrell, T. E., Shields, R. M., Allodi, M. A., Wood, E. K., Kirschner, K. N., Castonguay, T. C., Archer, K. A., and Shields, G. C.: Quantum Mechanical Study of Sulfuric Acid Hydration: Atmospheric Implications, J. Phys. Chem. A, 116, 2209–2224, https://doi.org/10.1021/jp2119026, 2012.

Temelso, B., Mabey, J. M., Kubota, T., Appiah-Padi, N., and Shields, G. C.: ArbAlign: A tool for optimal alignment of arbitrarily ordered isomers using the Kuhn–Munkres algorithm, J. Chem. Inf. Model., 57, 1045–1054, https://doi.org/10.1021/acs.jcim.6b00546, 2017.

Tikhonov, D. S., Gordiy, I., Iakovlev, D. A., Gorislav, A. A., Kalinin, M. A., Nikolenko, S. A., Malaskeevich, K. M., Yureva, K., Matsokin, N. A., and Schnell, M.: Harmonic Scale Factors of Fundamental Transitions for Dispersion-corrected Quantum Chemical Methods, ChemPhysChem, 25, e202400 547, https://doi.org/10.1002/cphc.202400547, 2024.

Torrie, G. M. and Valleau, J. P.: Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid, Chem. Phys. Lett., 28, 578–581, https://doi.org/10.1016/0009-2614(74)80109-0, 1974.

Tribello, G. A., Cuny, J., Eshet, H., and Parrinello, M.: Exploring the free energy surfaces of clusters using reconnaissance metadynamics, J. Chem. Phys., 135, 114 109, https://doi.org/10.1063/1.3628676, 2011.

Trolle, G. B., Kubečka, J., and Elm, J.: Modeling local aerosol surface environments: Clustering of pyruvic acid analogs, water, and Na^+, Cl^- ions, ACS Omega, 10, 1470–1485, https://doi.org/10.1021/acsomega.4c09196, 2025.

Tröstl, J., Chuang, W. K., Gordon, H., Heinritzi, M., Yan, C., Molteni, U., Ahlm, L., Frege, C., Bianchi, F., Wagner, R., and et al.: The role of low-volatility organic compounds in initial particle growth in the atmosphere, Nature, 533, 527–531, https://doi.org/10.1038/nature18271, 2016.

Vanhanen, J., Mikkilä, J., Lehtipalo, K., Sipilä, M., Manninen, H. E., Siivola, E., Petäjä, T., and Kulmala, M.: Particle size magnifier for nano-CN detection, Aerosol Sci. Technol., 45, 533–542, https://doi.org/10.1080/02786826.2010.547889, 2011.

Vehkamäki, H., Kulmala, M., Napari, I., Lehtinen, K. E. J., Timmreck, C., Noppel, M., and Laaksonen, A.: An improved parameterization for sulfuric acid–water nucleation rates for tropospheric and stratospheric conditions, J. Geophys. Res. Atmos., 107, AAC 3–1–AAC 3–10, https://doi.org/https://doi.org/10.1029/2002JD002184, 2002.

Vogt, E. and Kjaergaard, H. G.: Vibrational spectroscopy of the water dimer at jet-cooled and atmospheric temperatures, Annu. Rev. Phys. Chem., 73, 209–231, https://doi.org/annurev-physchem-082720-104659, 2022.

Wang, M., Kong, W., Marten, R., He, X.-C., Chen, D., Pfeifer, J., Heitto, A., Kontkanen, J., Dada, L., Kürten, A., Yli-Juuti, T., Manninen, H. E., Amanatidis, S., Amorim, A., Baalbaki, R., Baccarini, A., Bell, D. M., Bertozzi, B., Bräkling, S., Brilke, S., Murillo, L. C., Chiu, R., Chu, B., Menezes, L.-P. D., Duplissy, J., Finkenzeller, H., Carracedo, L. G., Granzin, M., Guida, R., Hansel, A., Hofbauer, V., Krechmer, J., Lehtipalo, K., Lamkaddam, H., Lampimäki, M., Lee, C. P., Makhmutov, V., Marie, G., Mathot, S., Mauldin, R. L., Mentler, B., Müller, T., Onnela, A., Partoll, E., Petäjä, T., Philippov, M., Pospisilova, V., Ranjithkumar, A., Rissanen, M., Rörup, B., Scholz, W., Shen, J., Simon, M., Sipilä, M., Steiner, G., Stolzenburg, D., Tham, Y. J., Tomé, A., Wagner, A. C., Wang, D. S., Wang, Y., Weber, S. K., Winkler, P. M., Wlasits, P. J., Wu, Y., Xiao, M., Ye, Q., Zauner-Wieczorek, M., Zhou, X., Volkamer, R., Riipinen, I., Dommen, J., Curtius, J., Baltensperger, U., Kulmala, M., Worsnop, D. R., Kirkby, J., Seinfeld, J. H., El-Haddad, I., Flagan, R. C., and Donahue, N. M.: Rapid growth of new atmospheric particles by nitric acid and ammonia condensation, Nature, 581, 184–189, https://doi.org/10.1038/s41586-020-2270-4, 2020.

Weber, K. H., Morales, F. J., and Tao, F.-M.: Theoretical study on the structure and stabilities of molecular clusters of oxalic acid with water, J. Phys. Chem. A, 116, 11 601–11 617, https://doi.org/10.1021/jp308499f, 2012.

Weber, K. H., Liu, Q., and Tao, F.-M.: Theoretical study on stable small clusters of oxalic acid with ammonia and water, J. Phys. Chem. A, 118, 1451–1468, https://doi.org/10.1021/jp4128226, 2014.

Weigend, F., Häser, M., Patzelt, H., and Ahlrichs, R.: RI-MP2: Optimized auxiliary basis sets and demonstration of efficiency, Chem. Phys. Lett., 294, 143–152, https://doi.org/10.1016/S0009-2614(98)00862-8, 1998.

Westermann, S., Langer, M., Boike, J., Heikenfeld, M., Peter, M., Etzelmüller, B., and Krinner, G.: Simulating the thermal regime and thaw processes of ice-rich permafrost ground with the land-surface model CryoGrid 3, Geosci. Model Dev., 9, 523–546, https://doi.org/10.5194/gmd-9-523-2016, 2016.

Woo, K. S., Chen, D. R., Pui, D. Y. H., and McMurry, P. H.: Measurement of Atlanta aerosol size distributions: Observations of ultrafine particle events, Aerosol Sci. Technol., 34, 75–87, https://doi.org/10.1080/02786820120056, 2001.

Woon, D. E. and Dunning, T. H.: Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon, J. Chem. Phys., 98, 1358–1371, https://doi.org/10.1063/1.464303, 1993.

Wu, H., Engsvang, M., Knattrup, Y., Kubečka, J., and Elm, J.: Improved configurational sampling protocol for large atmospheric molecular clusters, ACS Omega, 8, 45 065–45 077, https://doi.org/10.1021/acsomega.3c06794, 2023.

Wu, H., Knattrup, Y., Jensen, A. B., and Elm, J.: Cluster-to-particle transition in atmospheric nanoclusters, Aerosol Research, 2, 303–314, https://doi.org/10.5194/ar-2-303-2024, 2024.

Xu, C.-X., Jiang, S., Liu, Y.-R., Feng, Y.-J., Wang, Z.-H., Huang, T., Zhao, Y., Li, J., and Huang, W.: Formation of atmospheric molecular clusters of methanesulfonic acid–Diethylamine complex and its atmospheric significance, Atmos. Environ., 226, 117 404, https://doi.org/10.1016/j.atmosenv.2020.117404, 2020.

Xu, Y., Nadykto, A. B., Yu, F., Jiang, L., and Wang, W.: Formation and properties of hydrogen-bonded complexes of common organic oxalic acid with atmospheric nucleation precursors, J. Mol. Struct. THEOCHEM, 951, 28–33, https://doi.org/10.1016/j.theochem.2010.04.004, 2010.

Yazgi, D. and Olenius, T.: J-GAIN v1.1: A flexible tool to incorporate aerosol formation rates obtained by molecular models into large-scale models, Geosci. Model Dev., 16, 5237–5249, https://doi.org/10.5194/gmd-16-5237-2023, 2023.

Zapadinsky, E., Passananti, M., Myllys, N., Kurtén, T., and Vehkamäki, H.: Modeling on fragmentation of clusters inside a mass spectrometer, J. Phys. Chem. A., 123, 611–624, https://doi.org/10.1021/acs.jpca.8b10744, 2019.

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

Zhang, J. and Dolg, M.: ABCluster: The artificial bee colony algorithm for cluster global optimization, Phys. Chem. Chem. Phys., 17, 24 173–24 181, https://doi.org/10.1039/C5CP04060D, 2015.

Zhang, J. and Dolg, M.: Global optimization of clusters of rigid molecules using the artificial bee colony algorithm, Phys. Chem. Chem. Phys., 18, 3003–3010, https://doi.org/10.1039/C5CP06313B, 2016.

Zhang, R., Shen, J., Xie, H.-B., Chen, J., and Elm, J.: The role of organic acids in new particle formation from methanesulfonic acid and methylamine, Atmos. Chem. Phys., 22, 2639–2650, https://doi.org/10.5194/acp-22-2639-2022, 2022.

Zhao, Y. and Truhlar, D. G.: The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals, Theor. Chem. Acc., 120, 215–241, https://doi.org/10.1007/s00214-007-0310-x, 2007.

Zhu, Y.-P., Liu, Y.-R., Huang, T., Jiang, S., Xu, K.-M., Wen, H., Zhang, W.-J., and Huang, W.: Theoretical study of the hydration of atmospheric nucleation precursors with acetic acid, J. Phys. Chem. A, 118, 7959–7974, https://doi.org/10.1021/jp506226z, 2014.

Řezáč, J., Riley, K. E., and Hobza, P.: S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures, J. Chem. Theory Comput., 7, 2427–2438, https://doi.org/10.1021/ct2002946, 2011.