# A cluster-of-functional-groups approach for studying organic enhanced atmospheric cluster formation

Astrid Nørskov Pedersen, Yosef Knattrup, and Jonas Elm

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

**Correspondence:** Jonas Elm (jelm@chem.au.dk)

**Abstract.** The role of organic compounds in atmospheric new particle formation is difficult to disentangle due to the myriad of potentially important oxygenated organic molecules (OOMs) present in the atmosphere. Using state-of-the-art quantum chemical methods, we here employ a novel approach, denoted the "cluster-of-functional-groups" approach, for studying the involvement of OOMs in atmospheric cluster formation. Instead of the usual "trial-and-error" approach of testing the ability of experimentally identified OOMs to form stable clusters with other nucleation precursors, we here study which, and how many, intermolecular interactions that are required in a given OOM to form stable clusters. In this manner we can reverse engineer the elusive structure of OOM candidates that might be involved in organic enhanced atmospheric cluster formation.

We calculated the binding free energies of all combinations of donor/acceptor organic functional groups to investigate which functional groups that most preferentially bind with each other and with other nucleation precursors such as sulfuric acid and bases (ammonia, methyl-, dimethyl-, and trimethylamine). We find that multiple carboxyl groups leads to substantially more stable clusters compared to all other combinations of functional groups. Employing cluster dynamics simulations, we investigate how a hypothetically OOM composed of multiple carboxyl groups can stablize sulfuric acid – base clusters and provide recommendations for potential atmospheric multi-carboxylic acid tracer compounds that should be explicitly studied in the future.

The presented "cluster-of-functional-groups" approach is generally applicable and can be employed in many other applications, such as ion-induced nucleation and potentially in elucidating the structural patterns in molecules that facilitate ice nucleation.

## 1 Introduction

Secondary organic aerosols (SOA) constitute a major fraction of organic matter in the atmosphere (Jimenez et al., 2009) and influence both human health (Pelucchi et al., 2009) and global climate (Canadell et al., 2021). The term SOA formation usually comprises both the initial aerosol nucleation and the subsequent growth of existing aerosol particles via vapour uptake. However, while oxygentaed organic molecules (OOMs) are widely accepted as being important for aerosol growth, it remain ambiguous whether OOMs are important for aerosol nucleation in the atmosphere (Kirkby et al., 2023).

The onset of aerosol nucleation is governed by the formation of stable atmospheric molecular clusters (Kulmala et al., 2013). Sulfuric acid (Sipilä et al., 2010) and atmospheric bases such as ammonia (AM) (Kirkby et al., 2011) and amines (Almeida

et al., 2013; Jen et al., 2014; Glasoe et al., 2015) (methylamine (MA), dimethylamine (DMA), and trimetheylamine (TMA)) are important species for nucleation both over land and sea. In addition, iodine species have been shown to contribute to nucleation at coastal and marine environments (Baccarini et al., 2020; Beck et al., 2021; He et al., 2021, 2023). The global modelling study by Dunne et el. (Dunne et al., 2016) indicated that almost all nucleation occurring in the atmosphere involves sulfuric acid coupled with either ammonia and/or organics. However, the exact role of organics in aerosol nucleation remains elusive. Pure organic ion-induced nucleation has been reported both at the CLOUD chamber (Kirkby et al., 2016) and in the field (Rose et al., 2018). While some early work has inferred that organics play a role in multicomponent cluster formation (Zhang et al., 2004; Metzger et al., 2010; Schobesberger et al., 2013; Riccobono et al., 2014), it remains unknown whether this is also true for multicomponent SA–base–OOM nucleation or whether it proceeds via two isolated pathways (one for SA–base and one for pure OOMs). Both experiments (Lehtipalo et al., 2018; Kirkby et al., 2023) and quantum chemical calculations (Elm, 2019b) allude to a decoupled mechanism, but this has not yet been definitively confirmed.

The puzzle of the role of organics in aerosol nucleation originate from the fact that there exists a myriad of OOMs in the atmosphere. Volatile organic compounds (VOCs) are rapidly transformed into less volatile species in the air due to oxidation reactions with either OH, $NO_3$ or $O_3$, making it challenging to disentangle which OOMs are important for nucleation and which are important for SOA growth. Identification of OOMs is usually performed with mass spectrometer techniques such as the CI-APi-TOF (Jokinen et al., 2012). For instance, the seminal work by Ehn et al. (Ehn et al., 2014) measured the chemical composition of OOM monomers ($C_{10}H_{14-16}O_{7-11}$) and OOM covalently bound dimers ($C_{19-20}H_{28-32}O_{10-18}$) from $\alpha$-pinene ($C_{10}H_{16}$) oxidation. While this gives important information about the chemical composition of the OOMs, it yields no information about the exact functional groups or their arrangement in the molecules. Quantum chemical (QC) calculations can be applied to yield the cluster structures and in conjugation with cluster dynamics modelling can elucidate the role of specific OOMs in nucleation. However, this requires that the given OOM structure is known. This has led to an overabundance of QC studies, including work by our group, that investigate the binding affinity of different experimentally identified OOMs to themselves and other nucleation precursors. We recently reviewed the entire QC literature on the role of organics in cluster formation and unfortunately, not a single OOM has definitively been proven to participate in nucleation in the planetary boundary layer (Elm et al., 2023). The lack of progress could be ascribed to the fact that previous work has been looking at the wrong compounds. All studies have been performed on the organic monomers, while recent evidence from the CLOUD chamber has shown that it is in fact the covalently bound organic dimers, which have ultra low volatilities (ULVOCs) (Simon et al., 2020) that drive nucleation (Lehtipalo et al., 2018; Dada et al., 2023). We will hereon refer to these dimers as accretion products to clearly distinct them from dimer clusters. Studying large $C_{19-20}H_{28-32}O_{10-18}$ OOM accretion products such as those from $\alpha$-pinene oxidation is challenging using QC methods and to date not a single study exists. This is caused by the fact that QC methods scale steeply with system size. In addition, larger, more flexible molecules can exist in numerous conformations, which rapidly increases number of calculations required to identify the global minimum cluster structures.

Using state-of-the-art quantum chemical methods, we here employ a new approach, denoted the "cluster-of-functional-groups" approach (Elm et al., 2023), for studying the involvement of OOMs in atmospheric cluster formation. Instead of the usual "trial-and-error" approach of testing the ability of identified OOMs to form stable clusters with themselves or other nu-

cleation precursors, we here study exactly which and how many intermolecular interactions (in the form of functional groups) that are required in a given OOM to form stable clusters. This allows us to reverse engineer the potential structure of OOMs involved in organic enhanced atmospheric cluster formation. We explicitly study all possible donors (alcohol, peroxide) and acceptors (ether, epoxide, aldehyde, ketone, acid anhydride and ester), as well as carboxylic acids. Based on the functional groups that bind strongest, we extend the analysis to study the cluster dynamics of a hypothetical OOM binding to $(SA)_{1-2}(Base)_{1-2}$ clusters, where the bases are AM, MA, DMA and TMA.

## 2 Methods

### 2.1 Computational details

The semi-empirical GFN1-xTB (Grimme et al., 2017) energy calculations and geometry optimization were calculated with the xtb 6.4.0 program (Bannwarth et al., 2021). Gaussian16, version B.01 (Gau) was used for the Density Functional Theory (DFT) calculations with default convergence criteria. The $\omega$B97X-D functional (Chai and Martin, 2008) with the 6-31++G(d,p) basis set was chosen for geometry optimization and vibrational frequency calculations based on its performance in numerous benchmarks (Elm and Mikkelsen, 2014; Myllys et al., 2016; Elm and Kristensen, 2017; Schmitz and Elm, 2020; Jensen et al., 2022). Given that some vibrational frequencies were low, the quasi-harmonic approximation, as recommended by Grimme, (Grimme, 2012) was used to treat vibrational frequencies below $100$ cm$^{-1}$ as free rotors. ORCA 5.0.4 (Neese, 2012, 2018; Neese et al., 2020; Neese, 2022) was used to calculate single-point energy corrections using the Domain Local Pair Natural Orbital, DLPNO–CCSD($T_0$) (Riplinger and Neese, 2013; Riplinger et al., 2013) with the aug-cc-pVTZ basis set, using the TightSCF convergence criteria and the NormalPNO setting (Liakos et al., 2015). The workflow and subsequent data processing was automated using the JKCS program (Kubečka et al., 2023). All data have been added to the Atmospheric Cluster DataBase (ACDB) (Elm, 2019a). The atmospheric cluster dynamics were simulated using the Atmospheric Cluster Dynamics Code (ACDC) (McGrath et al., 2012) using a modified version of the code given in the online repository by Tinja Olenius (Olenius et al., 2013; Roldin et al., 2019) available in the JKCS suite. The simulations were done at 278.15 K with a constant coagulation sink of $-1.6 \times 10^{-3}$ s$^{-1}$, matching typical sink values (Dal Maso et al., 2008; Kontkanen et al., 2017). The clusters with a single sulfuric acid monomer were counted towards the total sulfuric acid concentration.

### 2.2 Configurational sampling

For the configurational sampling procedure, a funnel-type approach was employed (Temelso et al., 2018; Odbadrakh et al., 2020; Kubečka et al., 2019), where the level of theory is increased in each step as the number of candidate cluster structures is decreased. The workflow can be described as follows:

ABCluster $\rightarrow$ GFN1-xTB$^{OPT}$ $\rightarrow$ ArbAlign $\rightarrow$ $\omega$B97X-D$^{OPT+FREQ}$ $\rightarrow$ DLPNO$^{SP}$

ABCluster (Zhang and Dolg, 2015, 2016) was used to generate a large pool of cluster structures with the CHARMM forcefield. Settings as recommended by Kubečka et al. (Kubečka et al., 2019) with population size, $SN = 3000$, maximum generations,

$g_{\max} = 200$, and number of scout bees, $g_{\text{limit}} = 4$ was used. We employed both neutral, anionic, and cationic monomers in all combinations leading to overall neutral clusters and saved 1000 local minima for each combination. All local minima were optimized at the GFN1-xTB level. ArbAlign (Temelso et al., 2017) was used to filter out identical clusters, based on root-mean-square deviation (RMSD) values. Based on previous studies on atmospheric clusters, (Kildgaard et al., 2018b, a) an RMSD cutoff of 0.38 Å was chosen. As a large number of structures remained after ArbAlign, only the 100 cluster structures with the lowest electronic energy was selected for further optimization at the $\omega$B97X-D/6-31++G(d,p) level of theory. The five structures with the lowest free energy were then selected for calculation of single-point energies at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ-level.

## 2.3 The "cluster-of-functional-groups" approach

Small organic molecules were chosen to represent the functional groups that act as hydrogen bond donors (alcohol (CH$_3$OH) and peroxide (CH$_3$OOH)), as well as functional groups that act as hydrogen bond acceptors (ether (CH$_3$OCH$_3$), epoxide (C$_2$H$_4$O), aldehyde (CH$_3$CHO), ketone (CH$_3$COCH$_3$), acid anhydride (CH$_3$C(=O)OC(=O)CH$_3$) and ester (COOCH$_3$)). Besides these groups, carboxylic acid, (HCOOH), which is both an acceptor and a donor, was also included. The studied functional groups were included in the funnelling approach in order to identify the global minimum cluster structures. In this manner the functional groups should orient in the most favourable positions and can be viewed as ideal "contact points" for assembling the molecular backbone afterwards. Hence, the "cluster-of-functional-groups" should be viewed as a single proxy OOM. The strength of this approach is that we do not need to explicitly consider the origin of the OOM i.e anthropogenic (aromatics) or biogenic (isoprene/terpenes) or whether it is a monomer or an accretion product as the approach will inherently identify the structural patterns that are important directly based on their ability to participate in cluster formation (Elm et al., 2023). However, if the functional groups are optimally oriented far from each other, it could allude to that the potential compound is an accretion product.

We note the caveat that the "cluster-of-functional-groups" approach assumes that the binding free energies of the individual groups are additive. This might not necessarily be the case for realistic atmospheric OOMs, where several effects potentially make the free energies deviate from additivity and there can be expected some degree of cancellation of errors. For instance, the enthalpy contribution is expected to be more or less additive given that there are no intramolecular hydrogen bonds in the OOM. In addition, it is unlikely that the multiple moieties of any realistic OOM can simultaneously reach the ideal contact points without introducing some strain in the backbone, which will also lead to a higher enthalpy. The entropy contribution will most likely not be additive as a major contribution comes from the loss of translational and rotational degrees of freedom i.e. the conversion into lower-entropy vibrational degrees of freedom. For each clustering functional group six high-entropy degrees of freedom are lost, while only a total of six high-entropy degrees of freedom are lost in the OOM. Despite these deficiencies, the "cluster-of-functional-groups" approach can still be employed for screening purposes and to yield some indication of which combinations of functional groups might potentially be important in atmospheric cluster formation.

## 2.4 Atmospheric cluster dynamics

To study the time development of molecular cluster distributions, the free energy is used to solve the birth-death equations with ACDC (McGrath et al., 2012). The birth-death equations describe how clusters are destroyed and created by condensation and evaporation and are given as the change in concentration of cluster $i$,

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = \sum_{j<i} \beta_{j,(i-j)} c_j c_{(i-j)} + \sum_j \gamma_{(i+j)\to i} c_{(i+j)} - \sum_j \beta_{i,j} c_i c_j - \sum_{j<i} \gamma_{i\to j} c_i + Q_i - S_i. \tag{1}$$

Here, $j$ is another cluster in the system, $\beta_{i,j}$ is the collision coefficient between cluster $i$ and $j$, $\gamma_{i\to j}$ is the evaporation coefficient of cluster $i$ into a smaller cluster, of which one is cluster $j$. $Q_i$ covers outside sources of $i$, and $S_i$ other possible loss mechanisms for $i$. The collision coefficient is calculated using kinetic gas theory and the evaporation coefficient via mass balance based on calculated free energy,

$$\gamma_{(i+j)\to i} = \beta_{i,j} c_{\mathrm{ref}} \exp\left(\frac{\Delta G_{i+j} - \Delta G_i - \Delta G_j}{k_b T}\right), \tag{2}$$

where $c_{\mathrm{ref}}$ is the monomer concentration of the vapour used to calculate the free energies. The birth-death equations are generated by checking all possible cluster configurations and examining which evaporations and collisions are able to destroy or create a given cluster (McGrath et al., 2012). As it is not possible to explicitly simulate the entire cluster size range, a user-specific size limit has to be chosen where the clusters are considered stable against evaporation and are counted towards the formation rate. We will refer to these clusters as "outgrowing". For our systems, we chose the clusters with one additional acid compared to the cluster sizes we have data for [(SA)$_3$(base)$_2$, (SA)$_2$(OOM)$_2$(base)$_2$ and (SA)$_3$(OOM)$_1$(base)$_2$]. We exclude outgrowing collisions with bases as clusters composed of more bases than acids are usually not stable. The outgrowing clusters containing OOMs have one more molecule present compared to the pathway without OOM. This will slightly underestimate the contribution of OOM to the relative simulated $J_{\mathrm{potential}}$ rates. These outgrowing cluster sizes are still quite small and therefore artificially stabilize the systems as the critical cluster size is not necessarily captured well. To distinguish the calculated rates from actual nucleation rates, we will refer to the rates as "cluster formation potentials" ($J_{\mathrm{potential}}$). These rates illustrate the potential of the cluster to grow to larger sizes and corresponds to an upper-bound on the formation rate. We refer to the clusteromics I paper (Elm, 2021a) for additional information on "cluster formation potentials".

## 3 Results and discussion

### 3.1 Pure organic clusters

Initially, pure organic clusters were studied containing every possible donor-acceptor combination. This lead to a total of 21 dimer cluster structures. The calculated free energies, at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory, are given in Figure 1. The calculations are performed at 298.15 K and 1 atm.

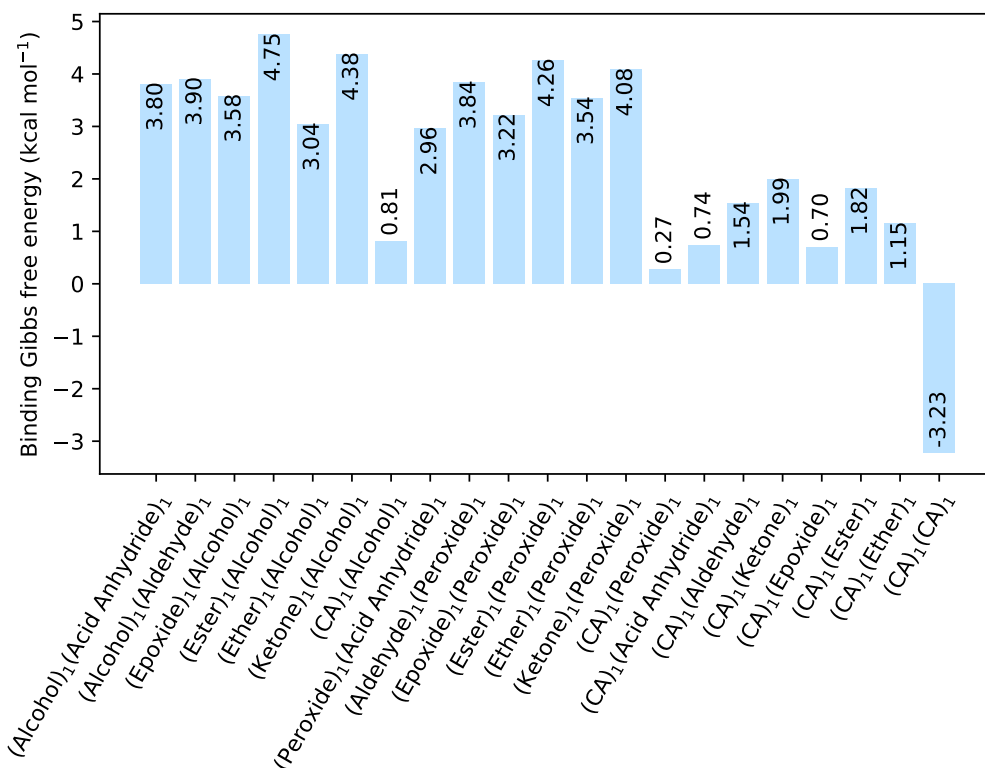



**Figure 1.** The binding free energy of all organic-organic interactions between all combinations of hydrogen bond donor and acceptor, calculated at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ //$\omega$B97X-D/6-31++G(d,p) level of theory with quasi-harmonic cutoff of 100 cm$^{-1}$, at 298.15 K and 1 atm. CA is short for Carboxylic Acid.

All the dimer clusters have a positive change in free energy, except for the dimer cluster with two carboxylic acid (CA) groups that has a $\Delta G = -3.23$ kcal mol$^{-1}$. This value is in excellent agreement with the experimental value of $-3.46$ kcal mol$^{-1}$ for the formic acid dimer cluster given by Vander et al. (Vander Auwera et al., 2007), at 296 K. However, this value corresponds to a high evaporation rate of the formic acid dimer. Thus, cluster formation involving only organic molecules is unlikely at 298.15 K and 1 atm, unless the molecule contains multiple carboxylic acid groups. In all cases the inclusion of a carboxyl group lowers the binding free energy considerably compared to the other functional groups. Hence, it is likely that other compounds besides organics are needed for forming the initial stable clusters.

## 3.2 SA–Organic clusters

As sulfuric acid is known to be essential for new particle formation, dimer clusters involving all the different organic functional groups and SA is studied. We studied the (SA)$_1$(organic)$_1$ dimer clusters where only a single organic functional group is present, as well as the (SA)$_1$(CA)$_1$(organic)$_1$ clusters that have been fully saturated by organic functional groups, meaning all donor and acceptor groups in the SA molecule are hydrogen-bonded. This led to a total of 46 clusters studied. The fully

165 saturated SA clusters all include at least one carboxylic acid group as well as two other identical organic groups, except for the case where there are two carboxylic acid groups in total as this fully saturates the SA molecule. The calculated free energies (at 298.15 K and 1 atm) are shown in Figure 2.

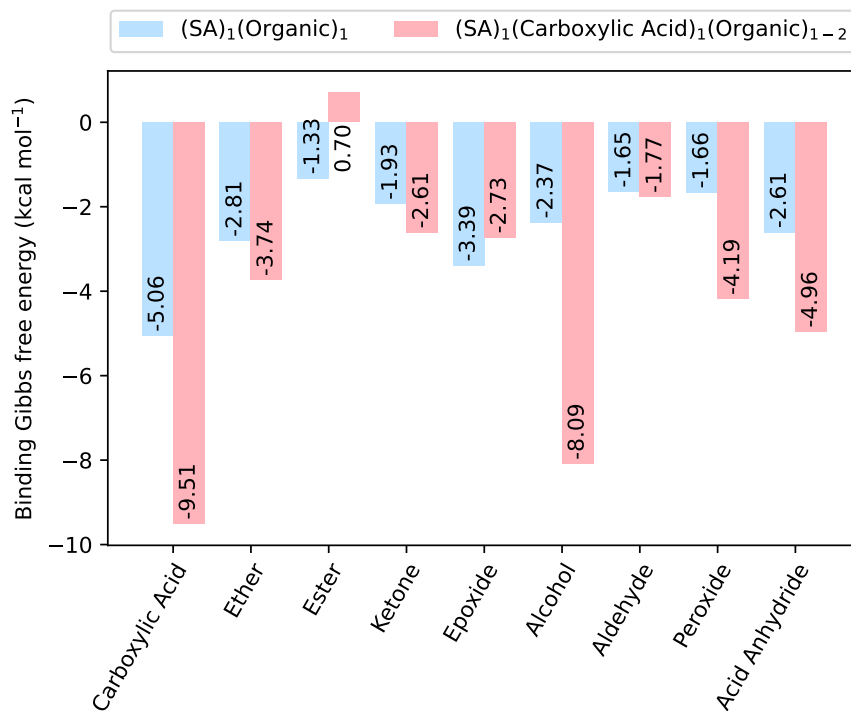



**Figure 2.** The binding free energy of clusters with one SA and one organic functional group, as well as with one SA fully saturated by organic functional groups, calculated at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory with quasi-harmonic cutoff of 100 cm$^{-1}$, at 298.15 K and 1 atm.

As expected, including SA decreases the free energy for all the dimer clusters when compared to the pure organic dimer
170 clusters. Generally, saturating the SA molecule with organic functional groups also results in a cluster lower in free energy. However, in cases where there is added two hydrogen bond donors to the (SA)$_1$(CA)$_1$ cluster, the addition of the second functional group leads to an increase in the binding free energy. For instance, this is the case for the clusters containing either two ester and two epoxide.

For the dimer clusters, (SA)$_1$(CA)$_1$ is the most stable with $\Delta G = -5.06$ kcal mol$^{-1}$, while for the clusters with a fully
175 saturated SA, it is (SA)$_1$(CA)$_2$, with $\Delta G = -9.51$ kcal mol$^{-1}$, indicating a trend of carboxylic acid being the most stabilizing functional group. This finding is consistent with the previous study by Elm et al. (Elm et al., 2017a). It should be noted that the (SA)$_1$(CA)$_1$(Alcohol)$_2$ cluster has a binding free energy very close to that of the (SA)$_1$(CA)$_2$ cluster, with a value of $\Delta G = -8.09$ kcal mol$^{-1}$. This suggests that multiple alcohol groups could also be important for cluster formation.

### 3.3 $(SA)_1(base)_1$–organic clusters

Next, bases are added to the clusters, to gauge its effect on the cluster stability. Initially, simple trimer clusters involving one organic functional group, one SA, and one base (either AM, MA, DMA or TMA) are studied. The calculated binding free energies of these clusters, at 298.15 K and 1 atm are presented in Figure 3.

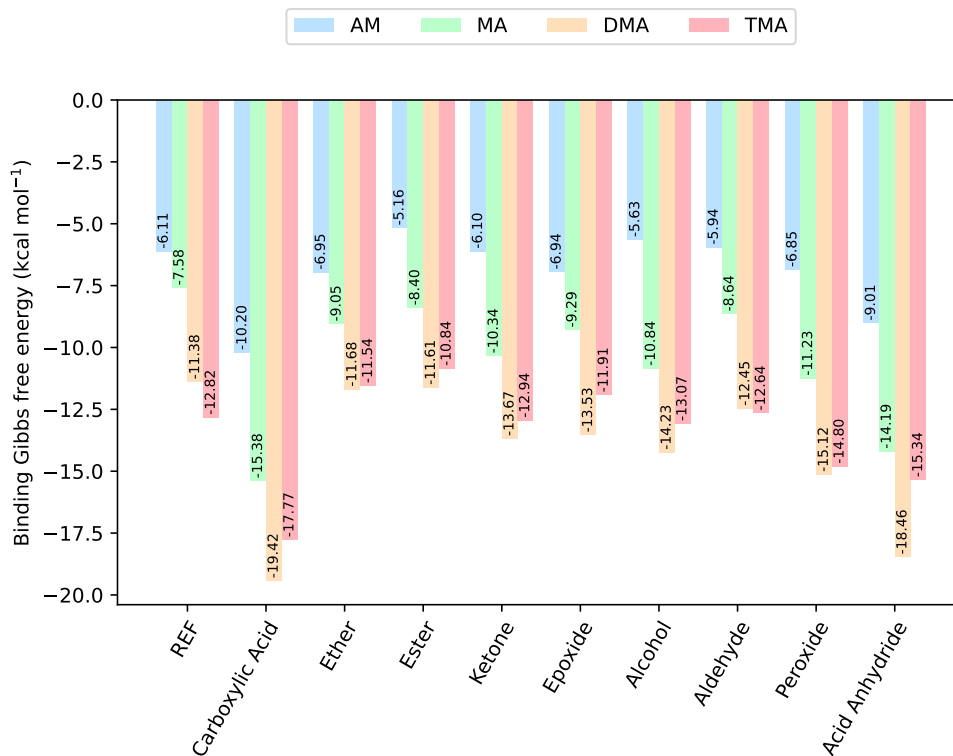



**Figure 3.** The binding free energy of the $(SA)_1(Base)_1(Organic)_1$ clusters, calculated at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ//ωB97X-D/6-31++G(d,p) level of theory with quasi-harmonic cutoff of 100 cm$^{-1}$, at 298.15 K and 1 atm. Carboxylic acid, the base, and sulfuric acid has been omitted in the labels for readability. REF is the reference $(SA)_1(Base)_1$ cluster without organics present.

In all cases the $(SA)_1(base)_1(organic)_1$ clusters are lower in free energy compared to the $(SA)_1(base)_1$ clusters. Hence, the addition of organics is thermodynamically favourable. For instance, the addition of a carboxylic acid group to the $(SA)_1(base)_1$ leads to a lowering of the free energy by $-4.09$, $-7.80$, $-8.09$ and $-4.95$ kcal mol$^{-1}$ for AM, MA, DMA and TMA, respectively.

Generally, the cluster binding free energies follow the basicity of the base with the following trend: DMA $\simeq$ TMA < MA < AM. This is consistent with previous work, where it has been shown both theoretically (Kurtén et al., 2008) and experimentally (Jen et al., 2014; Glasoe et al., 2015; Almeida et al., 2013) that amines are much more effective in driving NPF with SA than AM is, even when accounting for the large difference in atmospheric mixing ratios. The ordering of DMA and TMA depends

slightly on the system, with DMA dominantly leading to the most stable clusters. In particular, the $(SA)_1(DMA)_1(CA)_1$, $(SA)_1(TMA)_1(CA)_1$ and $(SA)_1(DMA)_1(Acid\ Anhydride)_1$ clusters have the lowest free energies, with $\Delta G$ of -19.42, -17.77 and -18.46 kcal mol$^{-1}$, respectively. Carboxylic acid groups are again found to form the most favourable interactions, with the cluster with acid anhydride only being about 1 kcal mol$^{-1}$ less stable.

To further explore which functional groups that could yield stable clusters, we added one more organic functional group to the $(SA)_1(Base)_1$ clusters. As carboxylic acid has been prevalent in forming the most stable clusters in the previous sections we added the constraint that one of the organics should be a carboxylic acid forming the $(SA)_1(Base)_1(CA)_1(Organic)_1$ clusters. The calculated binding free energies (at 298.15 K and 1 atm) are given in Figure 4.

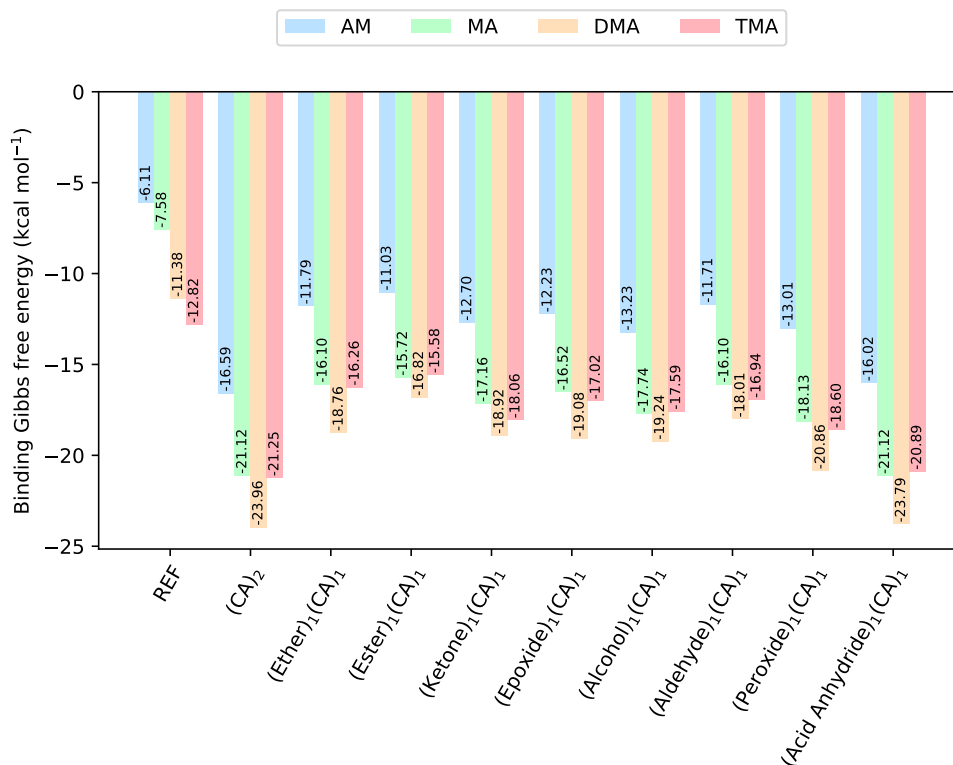



**Figure 4.** The binding free energy of the $(SA)_1(Base)_1(CA)_1(Organic)_1$ clusters, calculated at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory with quasi-harmonic cutoff of 100 cm$^{-1}$, at 298.15 K and 1 atm. CA is short for Carboxylic Acid. The base and sulfuric acid has been omitted in the labels for readability. REF is the reference $(SA)_1(Base)_1$ cluster without organics present.

The addition of one more organic functional group has a stabilizing effect on all the clusters. The cluster with two carboxylic acids is found to be lowest in free energy, with $\Delta G = -23.96$ kcal mol$^{-1}$ for the $(SA)_1(DMA)_1(CA)_2$ cluster. However, the cluster that contains an acid anhydride is also very close in free energy, with $\Delta G = -23.79$ kcal mol$^{-1}$. The cluster structures

can be seen in Figure 5, calculated at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory. In both
205   cases the organics can be seen to interact both with SA and the DMA molecules.

(SA)$_1$(DMA)$_1$(Acid Anhydride)$_1$(CA)$_1$        (SA)$_1$(DMA)$_1$(CA)$_2$

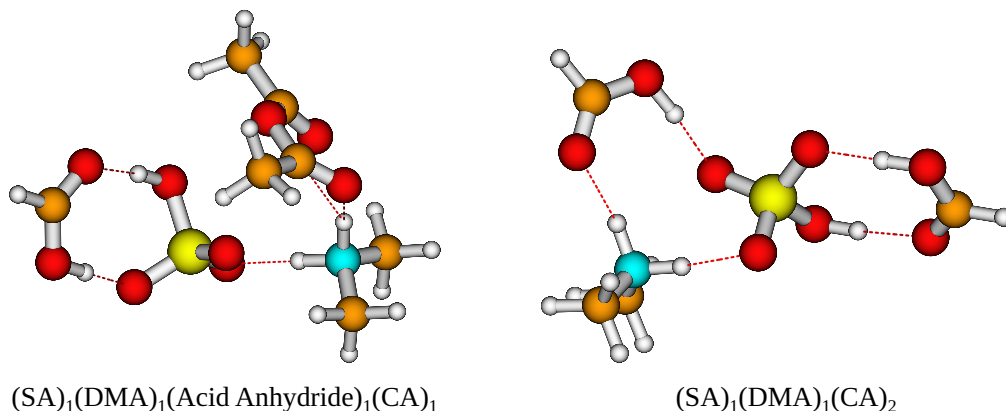



**Figure 5.** The (SA)$_1$(DMA)$_1$(Acid Anhydride)$_1$(CA)$_1$ and (SA)$_1$(DMA)$_1$(CA)$_2$ cluster geometries lowest in free energy. Calculated at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory with quasi-harmonic cutoff of 100 cm$^{-1}$, at 298.15 K and 1 atm. White = hydrogen, brown = carbon, red = oxygen, yellow = sulfur, blue = nitrogen. CA is short for Carboxylic Acid.

Furthermore, the (SA)$_1$(DMA)$_1$(Alcohol)$_1$(CA)$_1$, (SA)$_1$(DMA)$_1$(Aldehyde)$_1$(CA)$_1$, and (SA)$_1$(DMA)$_1$(Peroxide)$_1$(CA)$_1$
clusters also have low free energy values with $\Delta G$ of $-19.24$, $-18.01$ and $-20.86$ kcal mol$^{-1}$, respectively. Figure 6
presents the cluster structures. Again the organics are interacting with both SA and the DMA. Interestingly, the aldehyde
210   in the (SA)$_1$(DMA)$_1$(Aldehyde)$_1$(CA)$_1$ cluster preferentially binds to the SA and DMA compounds via several weak interac-
tions, instead of binding to the vacant S-OH group in SA. We found that the cluster where the aldehyde reside at the S-OH
group in SA is $0.62$ kcal mol$^{-1}$ higher in free than the presented structure in Figure 5 (See SI). Thus, the ability to form clusters
with organic molecules including multiple alcohols, aldehydes or peroxides as functional groups should be further studied in
the future.

215   (SA)$_1$(DMA)$_1$(Alcohol)$_1$(CA)$_1$        (SA)$_1$(DMA)$_1$(Aldehyde)$_1$(CA)$_1$        (SA)$_1$(DMA)$_1$(Peroxide)$_1$(CA)$_1$

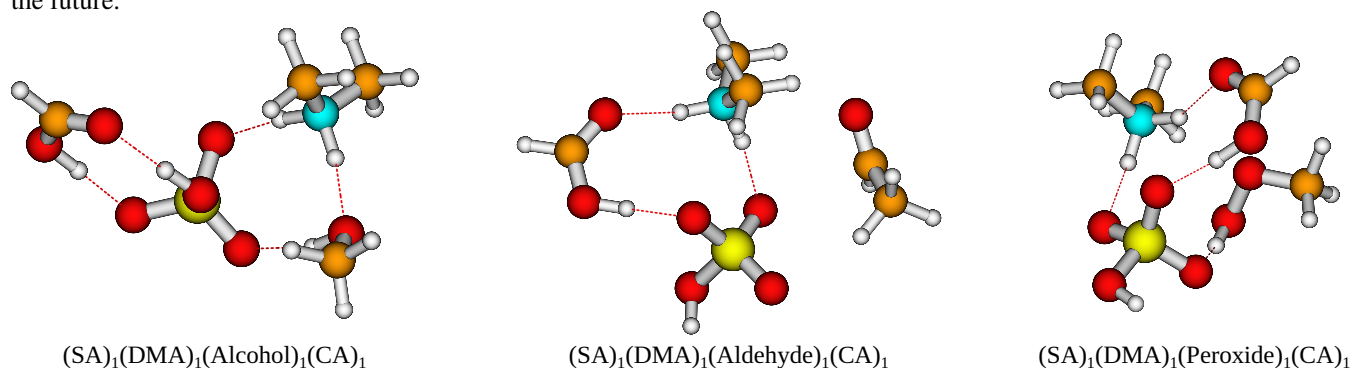



**Figure 6.** The (SA)$_1$(DMA)$_1$(Alchohol/Aldehyde/Peroxide)$_1$(CA)$_1$ cluster geometries lowest in free energy. Calculated at the DLPNO-CCSD(T$_0$)/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory with quasi-harmonic cutoff of 100 cm$^{-1}$, at 298.15 K and 1 atm. White = hydrogen, brown = carbon, red = oxygen, yellow = sulfur, blue = nitrogen. CA is short for Carboxylic Acid.

## 3.4 $(SA)_{1-2}(base)_{1-2}(CA)_3$ clusters

As the carboxylic acid groups consistently lead to the most stable clusters, the $(SA)_{1-2}(Base)_{1-2}(CA)_3$ clusters have also been studied. The calculated binding free energies, at 298.15 K and 1 atm are given in Figure 7.

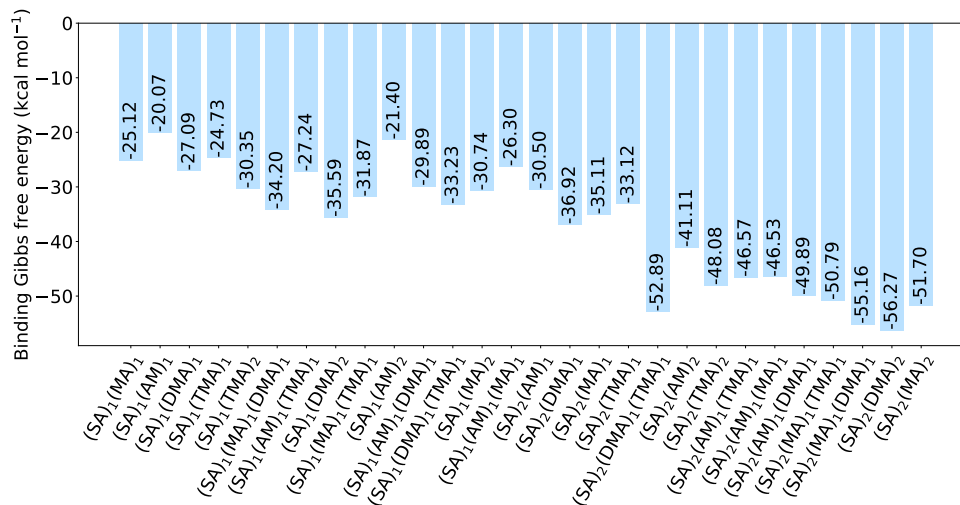



**Figure 7.** The binding free energy of the $(SA)_{1-2}(Base)_1(CA)_3$ clusters, calculated at the DLPNO-CCSD($T_0$)/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory with quasi-harmonic cutoff of 100 cm$^{-1}$, at 298.15 K and 1 atm. Carboxylic acid has been omitted in the labels for readability.

Logically, the inclusion of 3 carboxylic acid groups has a stabilizing effect, just as the inclusion of more SA and base does. Following the trend from earlier of DMA producing the most stable clusters, $(SA)_2(DMA)_2(CA)_3$ is the most stable cluster with $\Delta G = -56.27$ kcal mol$^{-1}$. Furthermore, DMA is found as a component in the three clusters that are lowest in free energy, with $(SA)_2(MA)_1(DMA)_1(CA)_3$ having $\Delta G = -55.16$ kcal mol$^{-1}$ and $(SA)_2(DMA)_1(TMA)_1(CA)_3$ having $\Delta G = -52.89$ kcal mol$^{-1}$. AM, on the other hand, is again the base that produces the clusters with the highest free energy, both alone and in combination with other bases. The two clusters highest in free energy are $(SA)_1(AM)_2(CA)_3$ with $\Delta G = -21.40$ kcal mol$^{-1}$ and $(SA)_1(AM)_1(CA)_3$ with $\Delta G = -20.07$ kcal mol$^{-1}$. Two SA molecules lead to more stable clusters than two DMA molecules, as $(SA)_2(DMA)_1(CA)_3$ has a slightly lower free energy compared to $(SA)_1(DMA)_2(CA)_3$ with a $\Delta G$ of -36.92 and -35.59 kcal mol$^{-1}$, respectively. While this is consistent with previous studies (Olenius et al., 2013; Elm, 2017), we see that the inclusion of the organic acids, decreases the difference between the two paths. The clusters lowest in free energy all have two SA molecules and two bases, as the inclusion of more SA and base gives rise to more acid-base interactions, as is illustrated in Figure 8.

$(SA)_1(DMA)_1(CA)_3$          $(SA)_2(DMA)_2(CA)_3$

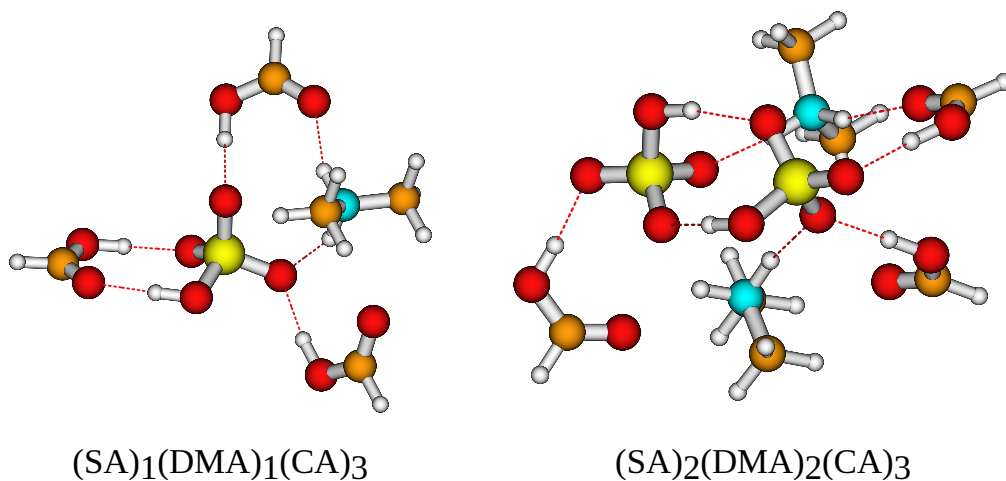
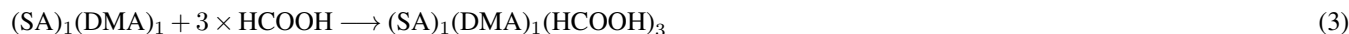



**Figure 8.** The $(SA)_{1-2}(DMA)_{1-2}(CA)_3$ cluster geometries lowest in free energy. Calculated at the DLPNO-CCSD($T_0$)/aug-cc-pVTZ//$\omega$B97X-D/6-31++G(d,p) level of theory with quasi-harmonic cutoff of 100 cm$^{-1}$, at 298.15 K and 1 atm. White = hydrogen, brown = carbon, red = oxygen, yellow = sulfur, blue = nitrogen. CA is short for Carboxylic Acid.

While the clusters become lower in free energy as they increase in size, it is not a guarantee that they are stable against evaporation in the atmosphere. For instance, the subsequent addition of more carboxylic acid groups to a the $(SA)_1(DMA)_1$ cluster leads to a decrease of $-8.04$, $-4.54$ and $-3.13$ kcal mol$^{-1}$ for the first, second and third addition, respectively. Hence, if these are treated as individual molecules it is unlikely that they form stable clusters at realistic atmospheric conditions. However, if all three functional groups are treated as single OOM, the free energy for adding a hypothetical idealized tricarboxylic to the $(SA)_1(DMA)_1$ cluster is $\Delta G = -15.71$ kcal mol$^{-1}$. It should be noted that this approach neglects the loss of high-entropy translation-and-rotation degrees of freedom (see discussion in Section 2.3). A pragmatic approach to remedy this effect would be to model the single OOM as a cluster of the three carboxylic acid groups. Essentially, we want to compare the following two reactions:

$$(SA)_1(DMA)_1 + 3 \times HCOOH \longrightarrow (SA)_1(DMA)_1(HCOOH)_3 \tag{3}$$

$$(SA)_1(DMA)_1 + (HCOOH)_3 \longrightarrow (SA)_1(DMA)_1(HCOOH)_3 \tag{4}$$

Calculating the binding free energy difference between reaction (3) and (4) we obtain:

$$\Delta\Delta G = \Delta G\Big[(HCOOH)_3\Big] - \Delta G\Big[3 \times HCOOH\Big] \tag{5}$$

Computing this value at 298.15 K and 1 atm leads to a perfect cancellation of errors with a $\Delta\Delta G$-value of -0.06 kcal mol$^{-1}$. However, as the enthalpy and entropy terms are different for reaction (3) and (4), we can expect a different temperature dependence. Recalculating $\Delta\Delta G$ at 278.15 K and 1 atm leads to +1.5 kcal mol$^{-1}$ difference. Hence, there is very little difference between the two methods of calculating the binding free energies and we will in the following stick to the simple method given by reaction (3). Nevertheless, this is a very strong binding and would correspond to a quite stable cluster. Hence, by employing the "cluster-of-functional-groups" approach we have identified that tricarboxylic acids are likely candidates for

forming stable clusters with SA and bases. This conclusion was already hypothesized, but not explicitly proven in our earlier work (Elm et al., 2017a).

### 3.5 Cluster formation potentials

255 Based on the calculated thermochemistry in the previous sections we can study how a hypothetical OOM composed of three carboxyl groups can potentially stablize sulfuric acid – base clusters. Letting the three carboxyl groups represent a single OOM we simulated the cluster formation potential ($J_{potential}$) of the $(SA)_{1-2}(base)_{1-2}(OOM)_1$ systems, with base = AM, MA, DMA and TMA. We calculated the collision coefficient of the OOM in the ACDC simulations based on the liquid density for formic acid and three times the mass of formic acid. To allow direct comparability, we conformed with our previous studies and used

260 the same vapour concentrations as in the clusteromics series of papers (Elm, 2021a, b, 2022; Knattrup and Elm, 2022; Ayoubi et al., 2023). Sulfuric acid was fixed at $10^6$ molecules cm$^{-3}$ and the concentration of the bases was studied giving two extremes with a "lower limit" and an "upper limit". These were set as follows: AM (10 ppt, 10 ppb), MA (1 ppt, 100 ppt) and DMA/TMA (1 ppt, 10 ppt). It should be noted that the low concentration limit likely best represents the actual concentrations observed in the ambient atmosphere.

265 The simulations were performed at 278.15 K and 1 atm using the settings described in Section 2.4. Figure 9 presents the simulated cluster formation potentials ($J_{potential}$) for the four bases as a function of OOM concentration (0-10 ppt).

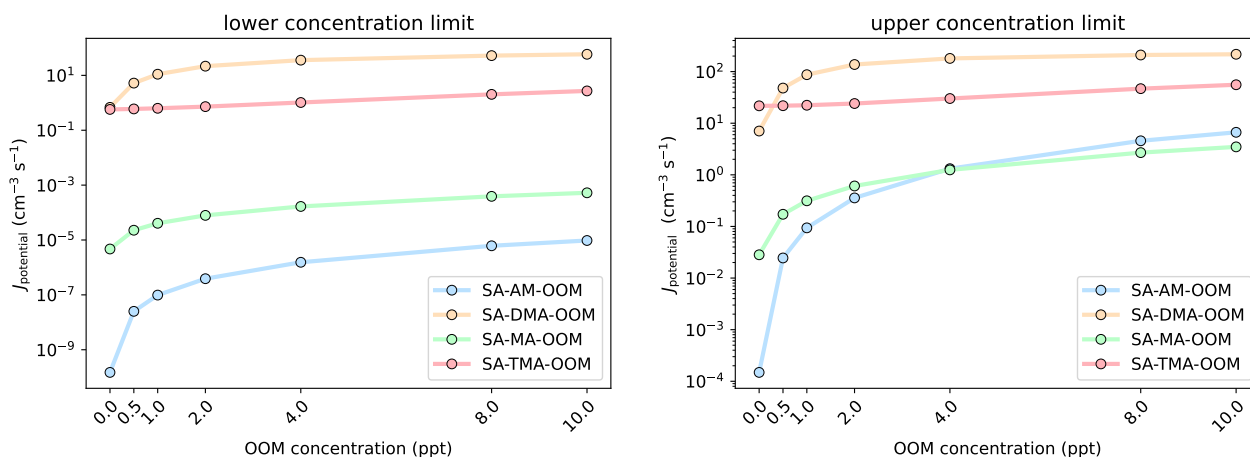



**Figure 9.** Simulated cluster formation potentials (in clusters cm$^{-3}$ s$^{-1}$) as a function of oxygenated organic molecule mixing ratio, in the lower concentration limit (*left*) and upper concentration limit (*right*). The simulations are performed at 278.15 K and 1 atm.

It is seen that the influence of the added OOM is highly dependent on the base in the SA–base–OOM systems. In the lower concentration limit the $J_{potential}$-value increases around one order of magnitude with 10 ppt OOM present for MA and

270 DMA. This is roughly four orders of magnitude for the SA–AM–OOM systems. However, the increase is from $1.49 \times 10^{-10}$ to $9.59 \times 10^{-6}$ cm$^{-3}$ s$^{-1}$, implying that the absolute cluster formation potential is still negligible. The cluster formation potential

of the SA–TMA–OOM system is found to be more or less unaffected by the presence of the OOM and thereby does not increase much with increased OOM concentration. This is caused by the three bulky methyl groups in TMA, which impede the attachment of the OOM due to steric hindrance (See SI). The formation potentials present the following trend at 10 ppt of OOM: DMA > TMA > MA > AM. In the upper concentration limit we see very similar trends, except that the SA–AM–OOM system begins to overtake the SA–MA–OOM system, due to the very high mixing ratio of 10 ppb of AM.

Inspecting the fluxes out of the systems, in the low concentration regime, show that the OOM (at 1 ppt) contribute to 97.96, 98.86, 72.44 and 7.52 % of the outgrowing fluxes for the AM, MA, DMA and TMA systems, respectively. The fluxes for each system at various concentrations can be seen in the SI. Hence, the more weakly bound the SA–base system is without thes OOM present, the more the OOM will contribute to the cluster formation potential. At 10 ppt of OOM 100 % of the outgrowing clusters contain an OOM for AM, MA and DMA, both in the lower and upper concentration regime of the bases. At 10 ppt of OOM the outgrowing clusters containing an OOM for TMA is 78.97 and 87.22 % in the lower and upper concentration limits. Again, it should be noted that the simulated $J_{\text{potential}}$-values show the potential to grow to larger sizes and not the actual nucleation rate. Hence, whether the OOM will evaporate from larger clusters or further contribtute to the cluster growth remains unknown.

Overall, we see that the hypothetical OOM substantially contributes to the cluster formation potential, but as this is an idealized compound, explicit SA–base–tricarboxylic acid clusters should be further studied. Cluster formation with 3-methyl-1,2,3-butanecarboxylic acid (MBTCA) (Müller et al., 2012) has previously been carried out for SA–MBTCA clusters (Ortega et al., 2016; Elm et al., 2017b) and SA–AM–MBTCA clusters (Myllys et al., 2017), but it remains ambiguous whether MBTCA is important for cluster formation. Recent evidence suggested that SA–MBTCA clusters primarily grow along the MBTCA coordinate, i.e. without the participation of SA (Elm, 2019b). Other potential atmospheric relevant tricarboxylic acids could be the 3-carboxyheptanoic acid (CHA) compound identified from limonene oxidation (Jaoui et al., 2006; Yasmeen et al., 2011). Recent work indicate that the CHA might be involved in SA–CHA cluster formation (Tan et al., 2022) and it would be worth to explicitly study SA–base–CHA clusters further. Finally, there has been identified a large accretion product, known as pinyl diaterpenylic ester (PDPE), from $\alpha$-pinene oxidation (Kristensen et al., 2013). Such a large compound has never been studied in atmospheric cluster formation and would be worth investigating in the future. This is also aligned with experimental observations, that deem large accretion products as the most likely OOMs to drive nucleation (Lehtipalo et al., 2018; Dada et al., 2023).

## 4  Conclusions

Using quantum chemical calculations we have employed a new method, denoted the "cluster-of-functional-groups" approach, for studying the role of oxygenated organic molecules (OOMs) in organic enhanced atmospheric cluster formation. Studying all combinations of organic donor/acceptor functional groups and their intermolecular interactions with sulfuric acid – base clusters it is found that carboyxyl groups leads to the most stable clusters. Based on our findings, we study the cluster formation potential of a hypothetical tricarboxylic acid molecule, represented as a cluster of three carboxyl groups. We find that the proxy

OOM composed of three carboxylic acid groups are present almost in all the relevant outgrowing clusters. Thereby the OOM is directly participating in SA–base-OOM cluster formation. Hence, the role of explicit tricarboxylic acids should further studied, and especially large flexible tricarboxylic acid accretion products might be important for cluster formation.

We hypothesize that a "cluster-of-functionals-groups" composed of multiple carbonyls, alcohols and hydroperoxide groups might rival the binding strength of three carboxylic acids in SA–base clustering and should be further studied in the future. The approach should also be extended to study ion-induced nucleation, which might subtantially enhance pure organic nucleation. In addition, the effect of temperature should be investigated, as all the functional groups will bind more strongly at lower temperatures.

Overall, our presented "cluster-of-functional-groups" approach is general and applicable to other topics where multifunctional compounds are important and where the exact chemical structure remains unknown. Hence, we speculate that the work could also be extended to study the potential molecular structures that can enhance ice nucleation in the form of identifying potential functional groups that facilitate the formation of hexagonal ice $I_h$ structures.

*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/pedersen24_org_interactions

*Author contributions.* Conceptualization: J.E.;
Methodology: A.N.P, Y.K., J.E.;
Formal analysis: A.N.P, Y.K.;
Investigation: A.N.P, Y.K.;
Resources: J.E.;
Writing - original draft: A.N.P, J.E.;
Writing - review & editing: A.N.P, Y.K., J.E.;
Visualization: A.N.P, Y.K.;
Project administration: J.E.;
Funding acquisition: J.E;
Supervision: J.E.

*Competing interests.* J.E is a member of the Editorial Board of Aerosol Research. The peer-review process was guided by an independent editor. The remaining authors have no conflict of interests to declare.

*Acknowledgements.* The authors thank the Independent Research Fund Denmark grant number 9064-00001B for financial support. This work was funded by the Danish National Research Foundation (DNRF172) through the Center of Excellence for Chemistry of Clouds.

The numerical results presented in this work were obtained at the Centre for Scientific Computing, Aarhus https://phys.au.dk/forskning/

335  faciliteter/cscaa/.

The authors thank Professor Merete Bilde for insightful discussions of the present work.

# References

Gaussian 16, Revision A.03, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., Gaussian, Inc., Wallingford CT, 2016.

340 Almeida, J., Schobesberger, S., Kürten, A., Ortega, I. K., Kupiainen-Määttä, O., Praplan, A. P., Adamov, A., Amorim, A., Bianchi, F., Breitenlechner, M., and et al.: Molecular Understanding of Sulphuric Acid-Amine Particle Nucleation in the Atmosphere, Nature, 502, 359–363, 2013.

Ayoubi, D., Knattrup, Y., and Elm, J.: Clusteromics V: Organic Enhanced Atmospheric Cluster Formation, ACS Omega, 8, 9621–9629, 2023.

Baccarini, A., Karlsson, L., Dommen, J., Duplessis, P., Vüllers, J., Brooks, I. M., Saiz-Lopez, A., Salter, M., Tjernström, M., Baltensperger,

345 U., Zieger, P., and Schmale, J.: Frequent New Particle Formation Over the High Arctic Pack Ice by Enhanced Iodine Emissions, Nat. Commun., 11, 4924, 2020.

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, 2021.

Beck, L. J., Sarnela, N., Junninen, H., Hoppe, C. J. M., Garmash, O., Bianchi, F., Riva, M., Rose, C., Peräkylä, O., Wimmer, D., et al.:

350 Differing Mechanisms of New Particle Formation at Two Arctic Sites, Geophys. Res. Lett., 48, 2021.

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cotrim da Cunha, L., Cox, P., Eliseev, A. V., Henson, S., Ishii, M., Jaccard, S., Koven, C., and et al.: Global Carbon and other Biogeochemical Cycles and Feedbacks, pp. 673–816, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021.

Chai, J. and Martin, H.: Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections, Phys. Chem.

355 Chem. Phys., 10, 6615–6620, 2008.

Dada, L., Stolzenburg, D., Simon, M., Fischer, L., Heinritzi, M., Wang, M., Xiao, M., Vogel, A. L., Ahonen, L., Amorim, A., et al.: Role of Sesquiterpenes in Biogenic New Particle Formation, Sci. Adv., 9, eadi5297, 2023.

Dal Maso, M., Hyvrinen, A., Komppula, M., Tunved, P., Kerminen, V.-M., Lihavainen, H., Viisanen, 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

360 Formation, Tellus B, 60, 495–508, 2008.

Dunne, E. M., Gordon, H., Kürten, A., Almeida, J., Duplissy, J., Williamson, C., Ortega, I. K., Pringle, K. J., Adamov, A., Baltensperger, U., et al.: Global Atmospheric Particle Formation from CERN CLOUD Measurements, Science, 354, 1119–1124, 2016.

Ehn, M., Thornton, J. A., Kleist, E., Sipilä, M., Junninen, H., Pullinen, I., Springer, M., Rubach, F., Tillmann, R., Lee, B., and et al.: A Large Source of Low-Volatility Secondary Organic Aerosol, Nature, 506, 476–479, 2014.

365 Elm, J.: Elucidating the Limiting Steps in Sulfuric Acid - Base New Particle Formation, J. Phys. Chem. A, 121, 8288–8295, 2017.

Elm, J.: An Atmospheric Cluster Database Consisting of Sulfuric Acid, Bases, Organics, and Water, ACS Omega, 4, 10 965–10 974, 2019a.

Elm, J.: Unexpected Growth Coordinate in Large Clusters Consisting of Sulfuric Acid and $C_8H_{12}O_6$ Tricarboxylic Acid, J. Phys. Chem. A, 123, 3170–3175, 2019b.

Elm, J.: Clusteromics I: Principles, Protocols and Applications to Sulfuric Acid - Base Cluster Formation, ACS Omega, 6, 7804–7814, 2021a.

370 Elm, J.: Clusteromics II: Methanesulfonic Acid-Base Cluster Formation, ACS Omega, 7, 17 035–17 044, 2021b.

Elm, J.: Clusteromics III: Acid Synergy in Sulfuric Acid-Methanesulfonic Acid-Base Cluster Formation, ACS Omega, 6, 15 206–15 214, 2022.

Elm, J. and Kristensen, K.: Basis Set Convergence of the Binding Energies of Strongly Hydrogen-Bonded Atmospheric Clusters, Phys. Chem. Chem. Phys, 19, 1122–1133, 2017.

375   Elm, J. and Mikkelsen, K. V.: Computational Approaches for Efficiently Modelling of Small Atmospheric Clusters, Chem. Phys. Lett., 615, 26–29, 2014.

Elm, J., Myllys, N., and Kurtén, T.: What is Required for Highly Oxidized Molecules to Form Clusters with Sulfuric Acid?, J. Phys. Chem. A., 121, 4578–4587, 2017a.

Elm, J., Myllys, N., Olenius, T., Halonen, R., Kurtén, T., and Vehkamäki, H.: Formation of Atmospheric Molecular Clusters Consisting of

380   Sulfuric Acid and $C_8H_{12}O_6$ Tricarboxylic Acid, Phys. Chem. Chem. Phys., 19, 4877–4886, 2017b.

Elm, J., Ayoubi, D., Engsvang, M., Jensen, A. B., Knattrup, Y., Kubečka, J., Bready, C. J., Fowler, V. R., Harold, S. E., Longsworth, O. M., and Shields, G. C.: Quantum Chemical Modeling of Organic Enhanced Atmospheric Nucleation: A Critical Review, WIREs Comput. Mol. Sci., 13, e1662, 2023.

Glasoe, W. A., Volz, K., Panta, B., Freshour, N., Bachman, R., Hanson, D. R., McMurry, P. H., and Jen, C.: Sulfuric Acid Nucleation: An

385   Experimental Study of the Effect of Seven Bases, J. Geophys. Res. Atmos., 120, 1933–1950, 2015.

Grimme, S.: Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory, Chem. Eur. J, 18, 9955–9964, 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.

390   Theory Comput., 13, 1989–2009, 2017.

He, X.-C., Tham, Y. J., Dada, L., Wang, M., Finkenzeller, H., Stolzenburg, D., Iyer, S., Simon, M., Kürten, A., Shen, J., et al.: Role of Iodine Oxoacids in Atmospheric Aerosol Nucleation, Science, 371, 589–595, 2021.

He, X.-C., Simon, M., Iyer, S., Xie, H.-B., Rörup, B., Shen, J., Finkenzeller, H., Stolzenburg, D., Zhang, R., and Baccarini, A.: Iodine Oxoacids Enhance Nucleation of Sulfuric Acid Particles in the Atmosphere, Science, 382, 1308–1314, 2023.

395   Jaoui, M., Corse, E., Kleindienst, T. E., Offenberg, J. H., Lewandowski, M., and Edney, E. O.: Analysis of Secondary Organic Aerosol Compounds from the Photooxidation of d-Limonene in the Presence of $NO_X$ and their Detection in Ambient PM2.5, Environ. Sci. Technol, 40, 3819–3828, 2006.

Jen, C. N., McMurry, P. H., and Hanson, D. R.: Stabilization of Sulfuric acid Dimers by Ammonia, Methylamine, Dimethylamine, and Trimethylamine, J Geophys. Res. Atmos., 119, 7502–7514, 2014.

400   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, 2022.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., and et al.: Evolution of Organic Aerosols in the Atmosphere, Science, 326, 1525–1529, 2009.

Jokinen, T., Sipilä, M., Junninen, H., Ehn, M., Lönn, G., Petäjä, J. H. T., Mauldin III, R. L., Kulmala, M., and Worsnop, D. R.: Atmospheric

405   Sulphuric Acid and Neutral Cluster Measurements using CI-APi-TOF, Atmos. Chem. Phys., 12, 4117–4125, 2012.

Kildgaard, J., Mikkelsen, K., Bilde, M., and Elm, J.: Hydration of Atmospheric Molecular Clusters II: Organic Acid - Water Clusters, J. Phys. Chem. A, 122, 8549–8556, 2018a.

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, 2018b.

410     Kirkby, J., Curtius, J., Almeida, J., Dunne, E., Duplissy, J., Ehrhart, S., Franchin, A., Gagne, S., Ickes, L., Kürten, and et al.: Role of Sulphuric Acid, Ammonia and Galactic Cosmic Rays in Atmospheric Aerosol Nucleation, Nature, 476, 429–433, 2011.

Kirkby, J., Duplissy, J., Sengupta, K., Frege, C., Gordon, H., Williamson, C., Heinritzi1, M., Simon, M., Yan, C., Almeida, J., et al.: Ion-induced Nucleation of Pure Biogenic Particles, Nature, 533, 521–525, 2016.

Kirkby, J., Amorim, A., Baltensperger, U., Carslaw, K. S., Christoudias, T., Curtius, J., Donahue, N. M., Haddad, I. E., Flagan, R. C., Gordon, 415     H., et al.: Atmospheric New Particle Formation from the CERN CLOUD Experiment, Nat. Geosci., 16, 948–957, 2023.

Knattrup, Y. and Elm, J.: Clusteromics IV: The Role of Nitric Acid in Atmospheric Cluster Formation, ACS Omega, 7, 31 551–31 560, 2022.

Kontkanen, J., Lehtipalo, K., Ahonen, L., Kangasluoma, J., Manninen, H. E., Hakala, J., Rose, C., Sellegri, K., Xiao, S., Wang, L., Qi, X., Nie, W., Ding, A., Yu, H., Lee, S., Kerminen, V.-M., Petäjää, T., and Kulmala, M.: Measurements of Sub-3 nm Particles using a Particle Size Magnifier in Different Environments: From Clean Mountain Top to Polluted Megacities, Atmos. Chem. Phys., 17, 2163–2187, 2017.

420     Kristensen, K., Enggrob, K. L., King, S. M., Worton, D. R., Platt, S. M., Mortensen, R., Rosenoern, T., Surratt, J. D., Bilde, M., Goldstein, A. H., and Glasius, M.: Formation and Occurrence of Dimer Esters of Pinene Oxidation Products in Atmospheric Aerosols, Atmos. Chem. Phys., 13, 3763–3776, 2013.

Kubečka, J., Besel, V., Kurtén, T., Myllys, N., and Vehkamäki, H.: Configurational Sampling of Noncovalent (Atmospheric) Molecular Clusters: Sulfuric Acid and Guanidine, J. Phys. Chem. A, 123, 6022–6033, 2019.

425     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, 2023.

Kulmala, M., Kontkanen, J., Junninen, H., Lehtipalo, K., Manninen, H. E., Nieminen, T., Petäjä, T., Sipilä, M., Schobesberger, S., Rantala, P., et al.: Direct Observations of Atmospheric Aerosol Nucleation, Science, 339, 943–946, 2013.

Kurtén, T., Loukonen, V., Vehkamäki, H., and Kulmala, M.: Amines are Likely to Enhance Neutral and Ion-induced Sulfuric Acid-water 430     Nucleation in the Atmosphere More Effectively than Ammonia, Atmos. Chem. Phys., 8, 4095–4103, 2008.

Lehtipalo, K., Yan, C., Dada, L., Bianchi, F., Xiao, M., Wagner, R., Stolzenburg, D., Ahonen, L. R., Amorim, A., Baccarini, A., et al.: Multicomponent New Particle Formation from Sulfuric Acid, Ammonia and Biogenic Vapors, Sci. Adv., 4, 1–9, 2018.

Liakos, D. G., Sparta, M., Kesharwani, M. K., Martin, J. M. L., and Neese, F.: Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory, J. Chem. Theory Comput., 11, 1525–1539, 2015.

435     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, 2012.

Metzger, A., Verheggen, B., Dommen, J., Duplissy, J., Prevot, A. S. H., Weingartner, E., Riipinen, I., Kulmala, M., Spracklen, D. V., Carslaw, K. S., and Baltensperger, U.: Evidence for the Role of Organics in Aerosol Particle Formation under Atmospheric Conditions, Proc. Natl. Acad. Sci. U.S.A., 107, 6646–6651, 2010.

440     Müller, L., Reinnig, M.-C., Naumann, K. H., Saathoff, H., Mentel, T. F., Donahue, N. M., and Hoffmann, T.: Formation of 3-methyl-1,2,3-butanetricarboxylic Acid via Gas Phase Oxidation of Pinonic Acid – A Mass Spectrometric Study of SOA Aging, Atmos. Chem. Phys., 12, 1483–1496, 2012.

Myllys, N., Elm, J., and Kurtén, T.: Density Functional Theory Basis Set Convergence of Sulfuric Acid-Containing Molecular Clusters, Comp. Theor. Chem., 1098, 1–12, 2016.

445     Myllys, N., Olenius, T., Kurtén, T., Vehkamäki, H., Riipinen, I., and Elm, J.: Effect of Bisulfate, Ammonia, and Ammonium on the Clustering of Organic Acids and Sulfuric Acid, J. Phys. Chem. A, 121, 4812–4824, 2017.

Neese, F.: The ORCA Program System, Wiley Interdiscip. Rev.: Comput. Mol. Sci., 2, 73–78, 2012.

Neese, F.: Software Update: The ORCA Program System, Version 4.0, Wiley Interdiscip. Rev.: Comput. Mol. Sci., 8, e1327, 2018.

Neese, F.: Software update: The ORCA program system—Version 5.0, WIREs Comput. Mol. Sci., 12, e1606, 2022.

450 Neese, F., Wennmohs, F., Becker, U., and Riplinger, C.: The ORCA Quantum Chemistry Program Package, J. Chem. Phys., 152, 224 108, 2020.

Odbadrakh, T. T., Gale, A. G., T., B. B., Temelso, B., and Shields, G. C.: Computation of Atmospheric Concentrations of Molecular Clusters from ab initio Thermochemistry, J. Vis. Exp., 8, e60 964, https://doi.org/10.3791/60964, 2020.

Olenius, T., Kupiainen-Määttä, O., Ortega, I. K., Kurtén, T., and Vehkamäki, H.: Free energy Barrier in the Growth of Sulfuric

455 Acid–Ammonia and Sulfuric Acid–Dimethylamine Clusters, J. Chem. Phys., 139, 084 312, https://doi.org/10.1063/1.4819024, 2013.

Ortega, I. K., Donahue, N. M., Kurtén, T., Kulmala, M., Focsa, C., and Vehkamäki, H.: Can Highly Oxidized Organics Contribute to Atmospheric New Particle Formation?, J. Phys. Chem. A, 120, 1452–1458, 2016.

Pelucchi, C., Negri, E., Gallus, S., Boffetta, P., Tramacere, I., and La Vecchia, C.: Long-term Particulate Matter Exposure and Mortality: A Review of European Epidemiological Studies, BMC Public Health, 9, 453, 2009.

460 Riccobono, F., Schobesberger, S., Scott, C. E., Dommen, J., Ortega, I. K., Rondo, L., Almeida, J., Amorim, A., Bianchi, F., Breitenlechner, M., and et al.: Oxidation Products of Biogenic Emissions Contribute to Nucleation of Atmospheric Particles, Science, 344, 717–721, 2014.

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, 2013.

Riplinger, C., Sandhoefer, B., Hansen, A., and Neese, F.: Natural Triple Excitations in Local Coupled Cluster Calculations With Pair Natural

465 Orbitals, J. Chem. Phys., 139, 134 101, 2013.

Roldin, P., Ehn, M., Kurtén, T., Olenius, T., Rissanen, M. P., Sarnela, N., Elm, J., Rantala, P., Hao, L., Hyttinen, N., and et al: The Role of Highly Oxygenated Organic Molecules in the Boreal Aerosol-cloud-climate System, Nat. Commun., 10, 4370, doi:10.1038/s41 467–019–12 338–8, 2019.

Rose, C., Zha, Q., Dada, L., Yan, C., Lehtipalo, K., Junninen, H., Mazon, S. B., Jokinen, T., Sarnela, N., and Sipilä", M.: Observations of

470 Biogenic Ion-induced Cluster Formation in the Atmosphere, Sci. Adv., 4, eaar5218, 2018.

Schmitz, G. and Elm, J.: Assessment of the DLPNO Binding Energies of Strongly Non-covalent Bonded Atmospheric Molecular Clusters, ACS Omega, 5, 7601–7612, 2020.

Schobesberger, S., Junninen, H., Bianchi, F., Lönn, G., Ehn, M., Lehtipalo, K., Dommen, J., Ehrhart, S., Ortega, I. K., Franchin, A., and et al.: Molecular Understanding of Atmospheric Particle Formation from Sulfuric Acid and Large Oxidized Organic Molecules, Proc. Natl.

475 Acad. Sci. U.S.A., 110, 17 223–17 228, 2013.

Simon, M., Dada, L., Heinritzi, M., Scholz, W., Stolzenburg, D., Fischer, L., Wagner, A. C., Kürten, A., Rörup, B., He, X.-C., and et al.: Molecular Understanding of New-particle Formation from $\alpha$-pinene between $-50$ and $+25$ °C, Atmos. Chem. Phys., 20, 9183–9207, 2020.

Sipilä, M., Berndt, T., Petäjä, T., Brus, D., Vanhanen, J., Stratmann, F., Patokoski, J., Mauldin, R. L., Hyvärinen, A.-P., Lihavainen, H., and

480 Kulmala, M.: The Role of Sulfuric Acid in Atmospheric Nucleation, Science, 327, 1243–1246, 2010.

Tan, S., Chen, X., and Yin, S.: Comparison Results of Eight Oxygenated Organic Molecules: Unexpected Contribution to New Particle Formation in the Atmosphere, Atmos. Environ., 268, 118 817, 2022.

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, 2017.

485    Temelso, B., Morrison, E. F., Speer, D. L., Cao, B. C., Appiah-Padi, N., Kim, G., and Shields, G. C.: Effect of Mixing Ammonia and Alkylamines on Sulfate Aerosol Formation, J. Phys. Chem. A, 122, 1612–1622, 2018.

Vander Auwera, J., Didriche, K., Perrin, A., and Keller, F.: Absolute Line Intensities for Formic Acid and Dissociation Constant of the Dimer, J. Chem. Phys, 126, 124 311, 2007.

Yasmeen, F., Szmigielski, R., Vermeylen, R., Gómez-González, Y., Surratt, J. D., Chan, A. W. H., Seinfeld, J. H., Maenhaut, W., and Claeys,
490    M.: Mass Spectrometric Characterization of Isomeric Terpenoic Acids From the Oxidation of $\alpha$-pinene, $\beta$-pinene, d-limonene, and $\Delta$3-carene in Fine Forest Aerosol, J. Mass Spectrom., 46, 425–442, 2011.

Zhang, J. and Dolg, M.: ABCluster: The Artificial Bee Colony Algorithm for Cluster Global Optimization, Phys. Chem. Chem. Phys., 17, 24 173–24 181, 2015.

Zhang, J. and Dolg, M.: Global Optimization of Clusters of Rigid Molecules Using the Artificial Bee Colony Algorithm, Phys. Chem. Chem.
495    Phys., 18, 3003–3010, 2016.

Zhang, R., Suh, I., Zhao, J., Zhang, D., Fortner, E. C., Tie, X., Molina, L. T., and Molina, M. J.: Atmospheric New Particle Formation Enhanced by Organic Acids, Science, 304, 1487–1490, 2004.