PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2017, **19**, 4877

Formation of atmospheric molecular clusters consisting of sulfuric acid and $C_8H_{12}O_6$ tricarboxylic acid†

Jonas Elm,*[a] Nanna Myllys,[a] Tinja Olenius,[b] Roope Halonen,[a] Theo Kurtén[c] and Hanna Vehkamäki[a]

Received 28th November 2016,
Accepted 10th January 2017

DOI: 10.1039/c6cp08127d

rsc.li/pccp

Using computational methods, we investigate the formation of atmospheric clusters consisting of sulfuric acid (SA) and 3-methyl-1,2,3-butanetricarboxylic acid (MBTCA), identified from α-pinene oxidation. The molecular structure of the clusters is obtained using three different DFT functionals (PW91, M06-2X and ωB97X-D) with the 6-31++G(d,p) basis set and the binding energies are calculated using a high level DLPNO-CCSD(T)/Def2-QZVPP method. The stability of the clusters is evaluated based on the calculated formation free energies. The interaction between MBTCA and sulfuric acid is found to be thermodynamically favourable and clusters consisting of 2–3 MBTCA and 2–3 SA molecules are found to be particularly stable. There is a large stabilization of the cluster when the amount of sulfuric acid–carboxylic acid hydrogen bonded interactions is maximized. The reaction free energies for forming the $(MBTCA)_{2-3}(SA)_{2-3}$ clusters are found to be similar in magnitude to those of the formation of the sulfuric acid–dimethylamine cluster. Using cluster kinetics calculations we identify that the growth of the clusters is essentially limited by a weak formation of the largest clusters studied, implying that other stabilizing vapours are required for stable cluster formation and growth.

1 Introduction

Atmospheric aerosols counteract the global warming caused by greenhouse gases, by scattering sun light back into space and by affecting the properties of clouds. A significant fraction of aerosol particles originates from new particle formation, but the chemical identity and relative significance of participating vapours remain highly uncertain. Experiments performed at the CLOUD chamber have shown that new particle formation from sulfuric acid and strong bases, such as dimethylamine, correlates well with atmospheric observations.[1] Significantly, less is known about the involvement of highly oxidized organic compounds in the initial steps of new particle formation. For instance Zhang *et al.* has demonstrated that new particle formation can be significantly enhanced by organic acids.[2]

Large quantities of volatile organic compounds (VOCs) are annually emitted into the ambient atmosphere from both anthropogenic and natural sources.[3] Terpenes, such as α-pinene, a monoterpene emitted from pine trees, constitute an important fraction of biogenic VOCs.[4,5] Terpenes are rapidly converted into highly oxidized species, in reactions initiated by the addition of either ozone or OH radicals to an endo/exocyclic double bond. Following the initial attack and rapid addition of O_2, the formation of highly oxidized terpene products can occur through an autoxidation mechanism. This involves intramolecular hydrogen shift reactions, followed by another addition of O_2 and eventually a termination reaction.[6–9] This type of mechanism has recently been uncovered for several proxies for monoterpene oxidation such as cyclohexene,[10] 1-methylcyclohexene and 4-methylcyclohexene,[11] and can account for the rapid formation of highly oxidized products with oxygen-to-carbon (O/C) ratios above 1. Due to the more complex molecular structure of terpenes, their autoxidation processes have not yet been fully resolved and the exact oxidation products remain elusive.[12]

Alternatively, terpenes can be oxidized through consecutive oxidation reactions. After the initial addition reaction, O_2 addition/rearrangements and termination, the product can be further oxidized by hydrogen abstraction reactions with OH radicals. Through several cycles of this consecutive oxidation mechanism a higher oxygen content can be incorporated into

[a] *Division of Atmospheric Sciences, Department of Physics, University of Helsinki, FI-00014 Helsinki, Finland. E-mail: jonas.elm@helsinki.fi*
[b] *Department of Environmental Science and Analytical Chemistry (ACES) and Bolin Centre for Climate Research, Stockholm University, 10691 Stockholm, Sweden*
[c] *Department of Chemistry, University of Helsinki, FI-00014 Helsinki, Finland*

† Electronic supplementary information (ESI) available: Calculated DFT enthalpies, entropies and Gibbs free energies. DLPNO single point energies. Calculated free energy surfaces, evaporation rates and ratio between collision rates and evaporation rates for all methods, at 258.15 K, 278.15 K and 298.15 K. Cartesian coordinates in *XYZ* format. See DOI: 10.1039/c6cp08127d

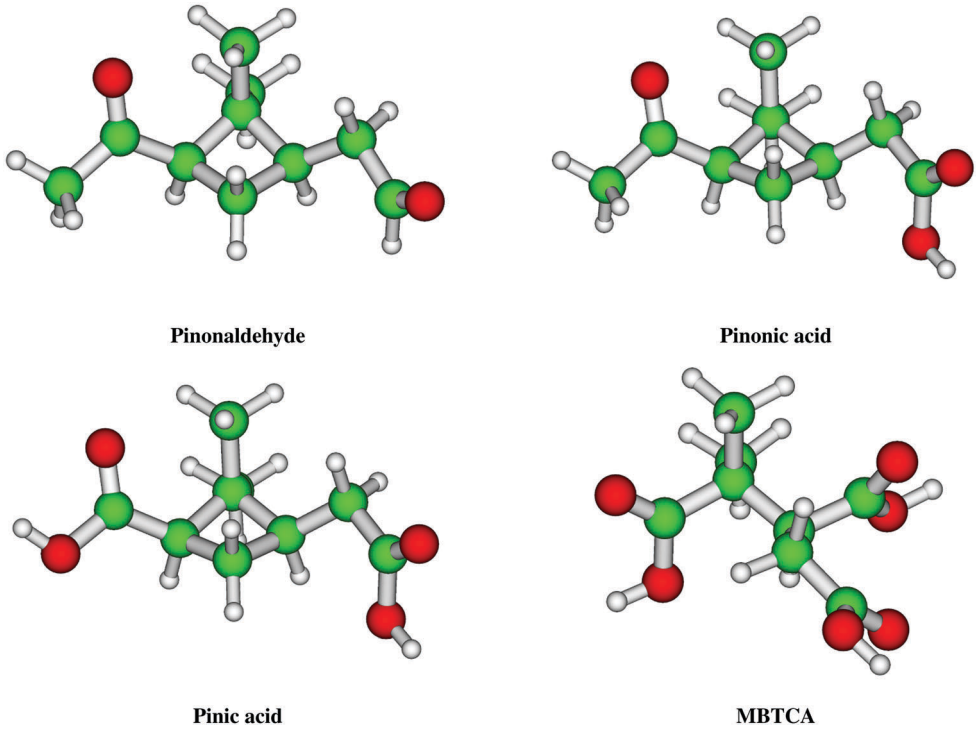

Fig. 1 The molecular structure of pinonaldehyde, pinonic acid, pinic acid and MBTCA. Color coding: green = carbon, red = oxygen and white = hydrogen.

the terpene precursor. From this process various distinct oxidation products of α-pinene have been identified, such as pinonaldehyde (O/C = 2/10), pinonic acid (O/C = 3/10) and pinic acid (O/C = 4/9),[13–17] as shown in Fig. 1. The consecutive oxidation is a slow process compared to autoxidation, as each generation of products requires another OH radical for initiating the process. At typical concentrations of the OH radical (in the order of 10^6 molecules cm^{-3} [18]) this implies that the reaction time for consecutive oxidation of monoterpene oxidation products, such as pinonaldehyde, is on the order of hours,[19] which makes it possible for the semi-volatile intermediate products to be removed *via* condensation on existing particles. It has recently been inferred that low-volatile monoterpene oxidation products with oxygen-to-carbon (O/C) ratios around and even above 1 are directly involved in the initial steps in atmospheric new particle formation.[20,21] The work by Schobesberger *et al.* indicates that 1–4 monoterpene oxidation products and 1–3 sulfuric acid molecules are required to obtain a stable cluster.[22] We recently studied the clustering potential of a $C_6H_8O_7$ autoxidation product from cyclohexene ozonolysis using computational methods.[23] We found that autoxidation products consisting mainly of peroxy acid and keto groups have a weak interaction both with themselves and with sulfuric acid. The participation of bases (ammonia and dimethylamine) did not promote the interaction enough to make $C_6H_8O_7$–sulfuric acid clustering occur under atmospheric conditions.[24] This suggests that the O/C ratio alone cannot be used directly as a metric for the volatility of a terpene oxidation product, thereby its potential to participate in new particle formation. This is further confirmed in the recent work by Kurtén *et al.* where no direct correlation was found between the O/C-ratio and the volatility of potential autoxidation products of α-pinene based on COSMO-RS calculations.[25] This illustrates that the specific number of strong hydrogen bonding groups, such as carboxylic acids, is very important for the ability to form clusters with sulfuric acid.

Using computational methods, we recently investigated the molecular interaction between pinic acid and sulfuric acid, and identified that the binding free energy was too weak to explain the formation of a stable cluster at atmospheric concentrations,[26] however the favourable reaction free energies indicated that pinic acid could contribute to the subsequent growth of an existing nucleus by condensation. This is further corroborated by the recent work by Depalma *et al.*, where a favourable reaction free energy of −16.8 kcal mol^{-1} was found, at the PW91/6-31++G(d,p) level of theory, for adding pinic acid to an existing $(H_2SO_4)_4(NH_3)_4$ cluster.[27] This indicates that for terpene oxidation products to be involved in the initial steps in new particle formation, more than two strong binding moieties are likely required, which would imply a higher oxygen content than that of pinic acid. Oxidation of pinonic acid by hydroxyl radicals can through complex pathways lead to the formation of 3-methyl-1,2,3-butanetricarboxylic acid (MBTCA) with an O/C-ratio of 6/8 (see Fig. 1).[28] Although the experimental formation mass yield of MBTCA has been identified to be a minuscule 0.61%, it could provide at least 10% of newly formed particle phase material.[29] Having three carboxylic acid moieties and O/C = 6/8, MBTCA represents one of the most likely candidates of an α-pinene oxidation product which could participate in atmospheric new particle formation.

Recently, Ortega *et al.* showed that MBTCA was indeed a more likely candidate for participating in atmospheric new particle formation compared to other α-pinene oxidation products such as pinonic and pinic acid.[30]

In this paper we wish to uncover the molecular interaction between MBTCA and sulfuric acid to identify the potential role of MBTCA in atmospheric new particle formation. Using density functional theory methods, we investigate some of the largest atmospheric molecular clusters explored to date, consisting of up to 3 MBTCA and 3 sulfuric acid molecules, reaching intermolecular distances of up to 1.7 nm. We further establish a mechanistic understanding of how monoterpene oxidation products are able to stabilize sulfuric acid clusters in the atmosphere. Using atmospheric cluster kinetics calculations we explore the limiting steps in the formation of new particles involving MBTCA and SA.

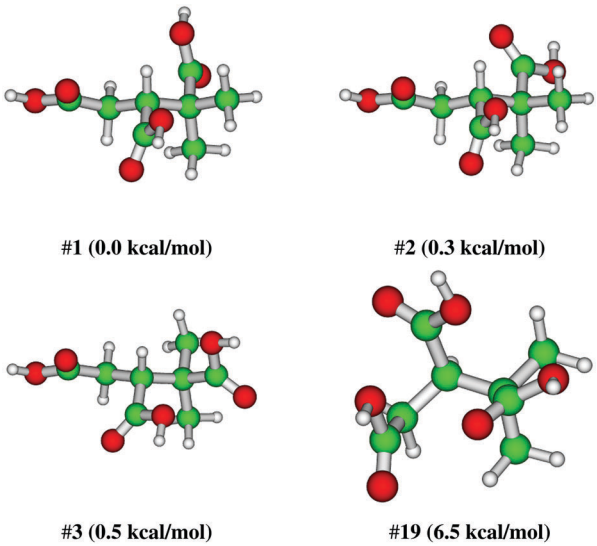

Fig. 2 The molecular structure of four different conformations of the MBTCA monomer, calculated at the M06-2X/6-31+G(d) level of theory. The relative stability of the conformers is shown in the brackets.

2 Methods

2.1 Computational details

All DFT calculations have been performed in Gaussian 09[31] using revision B.01, with default convergence criteria. We utilize the DFT functionals M06-2X, PW91 and ωB97X-D that have been identified to perform well in describing clusters of atmospheric relevance involving sulfuric acid.[32–35] Formation free energies were evaluated using harmonic oscillator and rigid rotor approximations, and unless otherwise specified are calculated at 298.15 K and 1 atm. For refining the electronic single point energies we use a domain local pair natural orbital coupled cluster (DLPNO-CCCD(T))[36,37] method as implemented in ORCA.[38]

2.2 Cluster formation

The MBTCA molecule has several internal rotational degrees of freedom and it is thereby a challenging task to identify the lowest Gibbs free energy conformation. Initially, all dihedral angles were scanned with a MMFF94 force field using the systematic rotor approach in Avogadro,[39] leading to more than 15 000 different conformations. The 100 conformations lowest in energy based on the force field calculation were stored and re-optimized using M06-2X/6-31+G(d). This led to 19 unique conformations with 13 conformations within 3 kcal mol^{-1} of the lowest identified one. Fig. 2 shows the molecular structures of conformations #1, #2, #3 and #19, and their relative stabilities in kcal mol^{-1}. All the structures close to the global minimum correspond to small rotations of the carboxylic acid groups, with the longest carbon backbone as linear as possible. In contrast, conformation #19 which was identified to be highest in free energy has a bent carbon backbone. All three tested DFT functionals predicted the same conformation to be the lowest in Gibbs free energy. The identification of such a large number of conformers explicitly shows the complexity of the potential energy surface of the system, even for the isolated monomer. To model the interaction with sulfuric acid, all 19 identified conformations were used as the starting point for forming the molecular clusters using the following semi-empirically guided technique:[40,41]

(1) In each cluster formation step 1000 randomly oriented molecules (lowest free energy H_2SO_4 or MBTCA) are randomly distributed around the target molecule/cluster.
(2) The structures are initially optimized using the semi-empirical PM6 method.
(3) For the converged structures a single-point M06-2X/6-31+G(d) energy is calculated.
(4) The structures are sorted based on the total energy and the dipole moment, and different conformations are identified.
(5) Conformations within 15 kcal mol^{-1} of the lowest identified conformation are geometry optimized and vibrational frequencies are calculated at the M06-2X/6-31+G(d) level.
(6) Remaining identified conformations within 3 kcal mol^{-1} of the lowest conformation are subsequently used for the next cluster formation step and the process is repeated from step 1.

By applying this systematic approach we should obtain a good estimate for the global minimum cluster structure. Due to the complexity of the potential energy surface we additionally included extensive manual sampling in each cluster formation step. Subsequently all conformations within 6 kcal mol^{-1} of the lowest identified conformation are further refined using the 6-31++G(d,p) basis set. The lowest identified conformations are then also geometry optimized and frequencies are calculated with the PW91 and ωB97X-D functionals using the 6-31++G(d,p) basis set. We recently showed that the reduction of the basis set from the large 6-311++G(3df,3pd) basis set to 6-31++G(d,p) had little effect on the thermal contribution to the Gibbs free energy (less than 1 kcal mol^{-1}), and did not change the subsequent calculation of the single point energy substantially.[42,43]

Using the smaller 6-31++G(d,p) basis set is necessary for the study at hand for allowing calculations up to the size of 1–3 MBTCA compounds and 1–4 sulfuric acid molecules without compromising the exploration of the configurational space. Unfortunately, the $(MBTCA)_3(H_2SO_4)_4$ cluster was still too computationally demanding and has been left out in the current study. The single point energy is subsequently calculated using DLPNO-CCSD(T) with a Def2-QZVPP basis set for each of the three geometries determined using the M06-2X, PW91 and ωB97X-D functionals. We recently showed that DLPNO-CCSD(T)/Def2-QZVPP yielded a consistent underestimation of the binding energy of sulfuric acid–base clusters and the values presented herein should thereby be considered as a lower bound for the formation free energies. The calculated binding energies were within 1 kcal mol^{-1} of the complete basis set limit, indicating that basis set superposition errors should not be a major source of error.[44] The final Gibbs free energies are presented as the average of all three values (DLPNO//DFT, with M06-2X, PW91 and ωB97X-D) and the scatter is reported as one standard deviation (σ). The σ-value thereby indicates the sensitivity of the Gibbs free energy to the applied quantum chemical method used in the geometry optimization and frequency calculation.

3 Results and discussion

From the configurational sampling routine numerous conformations were identified for each cluster. In Fig. 3 the structures of the lowest Gibbs free energy clusters are presented at the M06-2X/6-31+G(d) level of theory, with sulfuric acid simply abbreviated as SA. The corresponding reaction free energy of the different pathways for forming the clusters is shown in Fig. 4. The formation free energies are available for each functional in the ESI.† The simplest reaction between sulfuric acid and MBTCA is the formation of the (MBTCA)(SA) complex:

$$\mathrm{MBTCA + SA} \rightleftharpoons \mathrm{(MBTCA)(SA)}$$

Assuming mass-balance relations leads to the following complex concentration under equilibrium conditions:

$$[(\mathrm{MBTCA})(\mathrm{SA})] = [\mathrm{MBTCA}][\mathrm{SA}] \exp\left(\frac{-\Delta G}{k_B T}\right)$$

The atmospheric concentration of a given cluster is dependent on both the formation free energy and the concentration of the participating vapours. The formation of hydrogen bonds leads to a reduction in the enthalpy (ΔH). The clustering

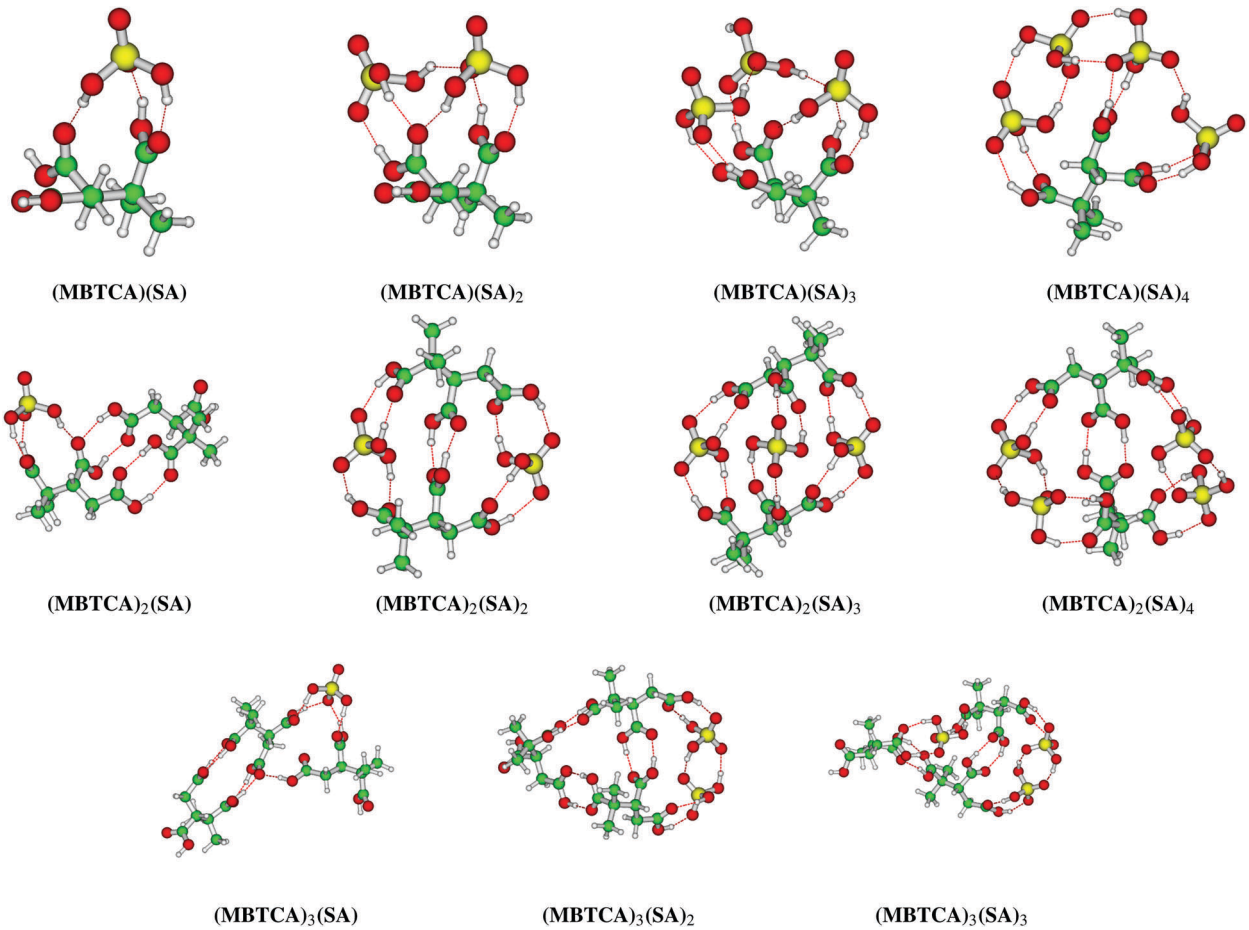

Fig. 3 The molecular structure of the identified clusters obtained at the M06-2X/6-31+G(d) level of theory. Color coding: green = carbon, yellow = sulfur, red = oxygen and white = hydrogen.

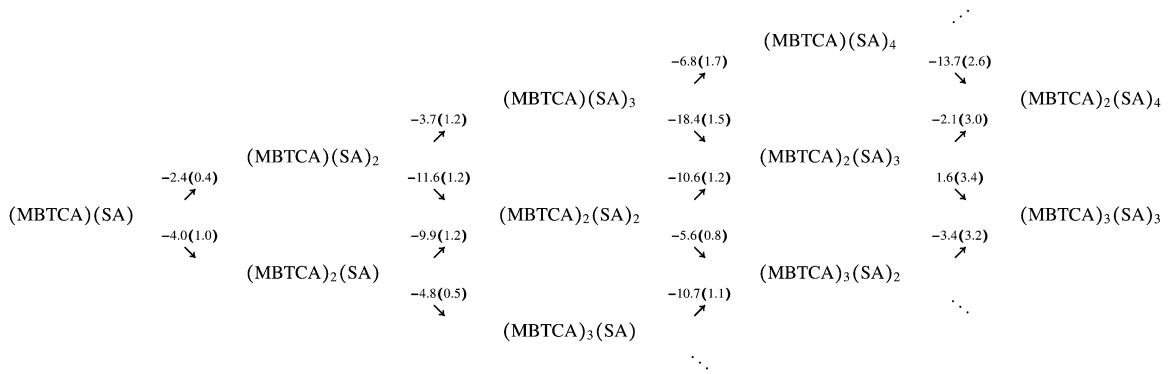

Fig. 4 Diagram for the studied cluster formation steps in the MBTCA–SA system. The presented values are the calculated reaction free energies (298.15 K and 1 atm), with the standard deviation (σ) shown in the parentheses. All values are reported in kcal mol^{-1}.

process is also accompanied by an entropy (ΔS) decrease,[45] as the formation of hydrogen bonds leads to a more constrained structure. The free energy is thereby dependent on these two opposite contributions, and hence even though more hydrogen bonds are formed in a given cluster, it does not necessarily lead to a more stable cluster; the specific types of interactions are equally important.

In Sections 3.1–3.7 we will only consider the reaction free energies of cluster formation (*i.e.* ΔG). For a cluster to be stable the collision rate that is equal to or higher than the evaporation rate is required. At typical atmospheric concentrations of sulfuric acid (mixing ratios of a few pptv) and pptv levels of MBTCA this would require a reaction free energy of at least −12 kcal mol^{-1} to yield a stable cluster.[26] We will refer to cluster stabilization as to how close the cluster formation reaction is to −12 kcal mol^{-1}. In Section 3.8 we will also take the effect of the vapour concentrations explicitly into account by performing cluster kinetics calculations.

3.1 Formation of the (MBTCA)(SA) cluster

The cluster formation between MBTCA and sulfuric acid involves the formation of three hydrogen bonds (see Fig. 3). The reaction free energy for forming this cluster is found to be −6.2 (σ = 0.2) kcal mol^{-1}. This process is slightly more favourable than the formation of the sulfuric acid dimer, with a reaction free energy of −5.2 (0.1) kcal mol^{-1} and the sulfuric acid ammonia cluster, with a reaction free energy of −4.6 (0.1) kcal mol^{-1}.[24] The formation of the (MBTCA)(SA) complex is, however, significantly less favourable than the formation of the sulfuric acid–dimethylamine cluster, where a reaction free energy value of −10.6 (0.9) kcal mol^{-1} has been reported.[24] The reaction free energy of the (MBTCA)$_2$ dimer is found to be −1.2 (0.8) kcal mol^{-1}, showing that the formation of the heterodimer between sulfuric acid and MBTCA is more favourable than the formation of either homodimers. Hanson and Lovejoy have reported the water mediated clustering of two sulfuric acid molecules.[46] Their experiments yielded a ΔH = −18.3 ± 1.8 kcal mol^{-1} and ΔS = −39.5 ± 7.8 kcal mol^{-1}. This corresponds to a ΔG-value of −6.5 kcal mol^{-1} for the formation of the sulfuric acid dimer. Our obtained value of −5.2 (0.1) kcal mol^{-1} is thereby slightly less favourable than the experimental results. This discrepancy can be attributed to the experimental value being effective values over the water distribution, while our calculations only take the dry clusters into account.

3.2 Addition of sulfuric acid molecules to (MBTCA)(SA)

The two subsequent additions of sulfuric acid molecules to the (MBTCA)(SA) complex are less favourable compared to the first, with values of −2.4 (0.4) and −3.7 (1.2) kcal mol^{-1}, respectively. This decrease is due to the sulfuric acid molecules ''competing'' to interact with the carboxylic acid moieties, leading to fewer direct sulfuric acid–carboxylic acid interactions, as seen in Fig. 3. More sulfuric acid–sulfuric acid interactions will thereby emerge in order to increase the amount of hydrogen bonds. The addition of the fourth sulfuric acid yields a higher stabilization, compared to the previous two, with a reaction free energy of −6.8 (1.7) kcal mol^{-1}. This can be attributed to the reduced strain in the MBTCA backbone, as the fourth sulfuric acid molecule allows direct bridging from one side of the MBTCA molecule to the other *via* an increased amount of sulfuric acid–sulfuric acid hydrogen bonded interactions.

The MBTCA–sulfuric acid cluster structures differ significantly from the previously studied pinic acid (PA)–sulfuric acid clusters.[26] The (MBTCA)(SA) complex forms three hydrogen bonds, whereas (PA)(SA) only is able to form two hydrogen bonds due to its two carboxylic acid groups being well separated by the rigid structure. Similarly, the bridging between the two sulfuric acid molecules seen in the (MBTCA)(SA)$_2$ cluster is not possible in the (PA)(SA)$_2$ cluster. In both the (MBTCA)(SA)$_3$ and (PA)(SA)$_3$ clusters a bridging between the sulfuric acid molecules is seen, but in the case of MBTCA it is possible to form more direct sulfuric acid–carboxylic acid hydrogen bonds due to its extra carboxylic acid group. This shows that the additional carboxylic acid group and flexible backbone of MBTCA lead to significantly different clusters compared to those previously studied.

3.3 Addition of MBTCA molecules to the (MBTCA)(SA) cluster

The first two additions of MBTCA molecules to the (MBTCA)(SA) complex are more favourable than the corresponding first two additions of sulfuric acid, with reaction free energies of −4.0 (1.0) and −4.8 (0.5) kcal mol^{-1}, respectively. From the molecular

structures of these clusters it is apparent that the higher stability is caused by the sulfuric acid molecule being capable of forming three hydrogen bonds to the MBTCA compound, while simultaneously the two MBTCA molecules can form four additional hydrogen bonds *via* direct carboxylic acid–carboxylic acid interactions. The $(MBTCA)_2(SA)$ cluster forms seven hydrogen bonds, while $(MBTCA)(SA)_2$ only forms six. We observe a similar pattern for $(MBTCA)_3(SA)$, which forms nine hydrogen bonds, while $(MBTCA)(SA)_3$ only forms eight.

3.4 Formation of the $(MBTCA)_2(SA)_2$ cluster

Adding a sulfuric acid molecule to the $(MBTCA)_2(SA)$ cluster leads to a favourable $(MBTCA)_2(SA)_2$ cluster with a reaction free energy of −9.9 (1.2) kcal mol^{-1}. This stability is similar in magnitude to the formation of the dimethylamine–sulfuric acid complex, which indicates that these clusters might be relatively stable against re-evaporation. The reaction free energy for forming the $(MBTCA)_2(SA)_2$ cluster is somewhat more favourable when adding an MBTCA molecule to the $(MBTCA)(SA)_2$ cluster, with a value of −11.6 (1.2) kcal mol^{-1}. The $(MBTCA)_2(SA)_2$ cluster can also be formed from the collision between two (MBTCA)(SA) pairs, with a reaction free energy of −7.7 kcal mol^{-1}. The particularly high stability of the $(MBTCA)_2(SA)_2$ cluster can be attributed to the formation of ten hydrogen bonds, with eight of them involving direct sulfuric acid–carboxylic acid interactions, and the remaining two from interaction between two carboxylic acid groups in MBTCA.

3.5 Formation of $(MBTCA)_3(SA)_2$

Adding a sulfuric acid molecule to the $(MBTCA)_3(SA)$ cluster leads to a high stabilization with a reaction free energy of −10.7 (1.1) kcal mol^{-1}. The formation of the $(MBTCA)_3(SA)_2$ cluster by adding an MBTCA molecule to the $(MBTCA)_2(SA)_2$ cluster is less favourable with a reaction free energy of −5.6 (0.8) kcal mol^{-1}. This lower stabilization is due to the reactant $(MBTCA)_2(SA)_2$ cluster already being highly stable. The $(MBTCA)_3(SA)_2$ cluster has a total of 12 hydrogen bonded interactions and due to steric effects it is not possible to identify a cluster where the amount of sulfuric acid–carboxylic acid interactions is maximized without involving the formation of less hydrogen bonds.

3.6 Formation of $(MBTCA)_2(SA)_3$

Both pathways for forming the $(MBTCA)_2(SA)_3$ cluster by monomer addition are highly favourable. The addition of a sulfuric acid molecule to the $(MBTCA)_2(SA)_2$ cluster yields a reaction free energy of −10.6 (1.2) kcal mol^{-1}. The formation by adding a MBTCA molecule to the $(MBTCA)(SA)_3$ cluster is very favourable, with a reaction free energy of −18.4 (1.5) kcal mol^{-1}, originating from the instability of the reactant $(MBTCA)(SA)_3$ cluster. The high stability of the $(MBTCA)_2(SA)_3$ cluster originates from the fact that 12 direct sulfuric acid–carboxylic acid hydrogen bonded interactions are present.

3.7 Formation of $(MBTCA)_2(SA)_4$ and $(MBTCA)_3(SA)_3$

The addition of sulfuric acid molecules to $(MBTCA)_2(SA)_3$ and $(MBTCA)_3(SA)_2$ clusters yields high reaction free energies, due to the high stability of the reactant clusters. The high stability of the $(MBTCA)_2(SA)_2$, $(MBTCA)_2(SA)_3$ and $(MBTCA)_3(SA)_2$ clusters can be understood from the molecular structures shown in Fig. 3, which shows that in general there is a high stabilization when the amount of direct sulfuric acid–carboxylic acid hydrogen bonded interactions is maximized. The higher formation free energies for forming larger clusters with more than three sulfuric acid and three MBTCA molecules could indicate that more sulfuric acid molecules are required to allow the maximum amount of sulfuric acid–carboxylic acid interactions to be reached. As three MBTCA molecules have a total of nine carboxylic acid groups, it can be assumed that 4–5 sulfuric acid molecules are required to maximize the amount of direct interactions.

3.8 $(MBTCA)_a(SA)_b$ cluster kinetics

From the law of mass action, we can obtain the actual Gibbs free energy surface of the clusters at a given concentration of sulfuric acid and MBTCA. The calculations were performed at a sulfuric acid concentration of 10^7 molecules cm^{-3}, with 1 pptv of MBTCA, using the DLPNO-CCSD(T)/Def2-QZVPP//DFT/6-31++G(d,p) or the pure DFT/6-31++G(d,p) quantum chemical formation free energies as input. Fig. 5 shows the actual formation free energy surface (in kcal mol^{-1}) on the MBTCA–SA grid, at 298.15 K. While the absolute value of the formation free energies varies between the different methods, a similar trend is seen for all four cases, with a free energy barrier for forming larger clusters. For any given cluster there is no growth direction (*i.e.* addition of either MBTCA or SA) that leads to a lower formation free energy. Following the lowest free energy path, based on the DLPNO surface, the cluster formation is initiated by the formation of (MBTCA)(SA). From this cluster the formation of the $(MBTCA)_2(SA)$ cluster has a lower free energy barrier than forming the $(MBTCA)(SA)_2$ cluster. This is in agreement with the recent computational study by Ortega *et al.*[30] and the experimental evidence from the CLOUD chamber.[21,22] In $(MBTCA)_2(SA)$ the most probable growth path is the addition of sulfuric acid molecules to yield the $(MBTCA)_2(SA)_2$ and $(MBTCA)_2(SA)_3$ clusters. Thus, the most favourable path is not along the MBTCA–SA diagonal. In general it is seen that the free energy steeply increases towards the system boundaries, which implies that the growth within the system is unfavorable, and cluster formation is suppressed. No critical cluster exists within the simulation box, and the steepness of the increase indicates that at these concentrations either a critical cluster does not exist at all, or will be at least the size of the boundary clusters.

The barriers are significantly higher for the DLPNO free energy surface than for the DFT surface. It has previously been established that the DLPNO/Def2-QZVPP binding energies are underestimations compared to canonical Coupled Cluster theory.[44] M06-2X/6-31++G(d,p) shows the lowest barriers in the studied system. However, using M06-2X with a low basis set is known to overestimate the binding energies.[23,26] This implies that the DLPNO and M06-2X results should represent a lower and upper bound, respectively, for the free energy barriers.

The calculated formation free energies are dependent on the temperature. Cluster formation is associated with a negative

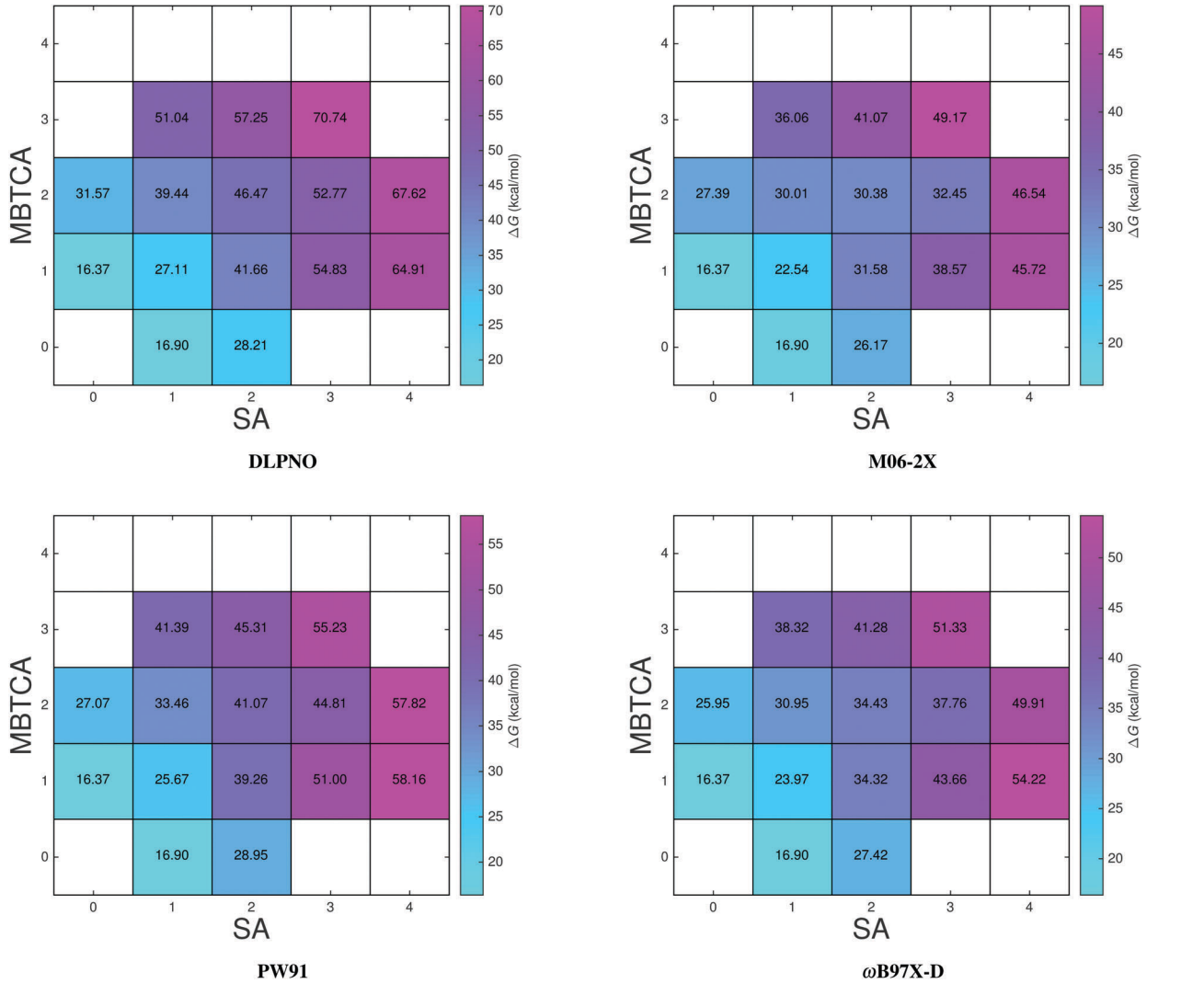

Fig. 5 Free energy surface (kcal mol^{-1}) of the MBTCA–SA system calculated with either DLPNO, M06-2X, PW91 or ωB97X-D. [H_2SO_4] = 10^7 molecules cm^{-3} and [MBTCA] = 1 pptv. Calculations were performed at 298.15 K.

change in the entropy, and the formation free energies will thereby be more favourable at lower temperatures. The Gibbs free energy surfaces have also been calculated at 278.15 K and 258.15 K (see the ESI†). The free energy barriers are significantly reduced at lower temperature. Each decrease of 20 K yields a lowering of $\sim$10 kcal mol^{-1} in the free energy barriers, for the growth of the largest clusters.

As the formation free energy increases towards the system boundaries, new particles are not formed in the present system. To form new particles the collision rate of monomers to the clusters must exceed the cluster evaporation rates, beyond some cluster size. By investigating the ratio between the sulfuric acid monomer collision rates and the total evaporation rate ($\beta_{SA}C_{SA}/\sum\gamma$) it can be estimated whether the clusters will grow under atmospheric conditions (see Fig. 6). Here $\sum\gamma$ is the sum of the individual evaporation rates for a given cluster. The calculations were performed using the DLPNO-CCSD(T)/Def2-QZVPP//DFT/6-31++G(d,p) or the pure DFT/6-31++G(d,p) formation free energies as input. A sulfuric acid concentration of 10^7 molecules cm^{-3} was used, with 1 pptv of MBTCA at 278.15 K.

For all the cases the ratio is below 1, indicating that the clusters will evaporate faster than they collide with sulfuric acid monomers. The corresponding calculations with collision of MBTCA molecules show similar trends and can be seen in the ESI.† Regardless of the method, the $(MBTCA)_2(SA)_2$ and $(MBTCA)_2(SA)_3$ clusters are more stable against evaporation than all other clusters. This indicates that in order to reach a stable cluster, at least 8–12 direct sulfuric acid–carboxylic acid interactions are required. While these clusters cannot grow by further addition of sulfuric acid or MBTCA, they could act as seeds for further addition of other vapour molecules. The complex pathways for forming MBTCA from oxidation of pinonic acid make it improbable that MBTCA concentrations exceed 1–10 pptv. New particle formation events are predominant during the spring and summer, meaning that temperatures below 278.15 K in the lower boundary layer are also unlikely. It is thereby highly unlikely that MBTCA and sulfuric acid by

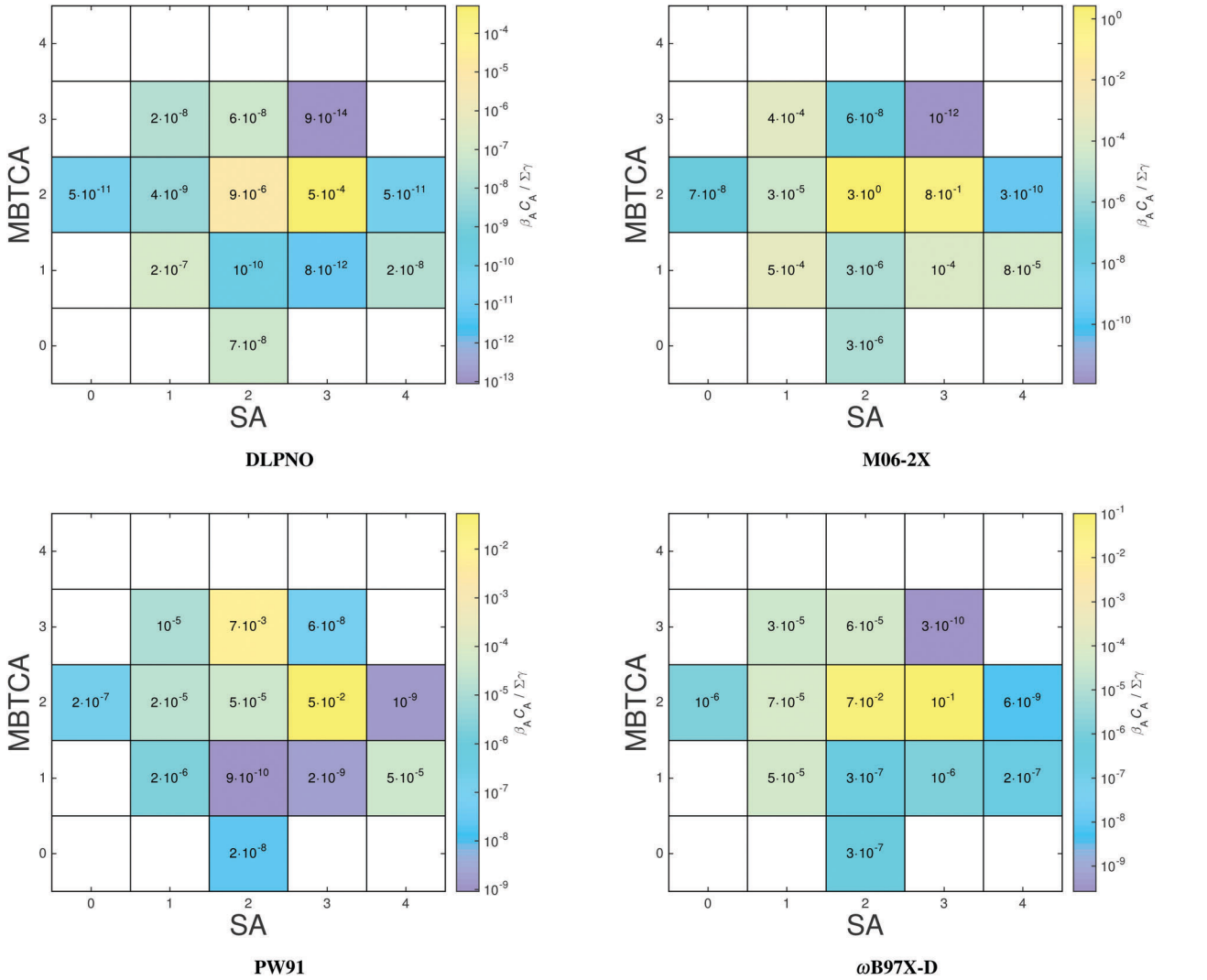

Fig. 6 The ratio of the rate of collisions with sulfuric acid molecules to the total evaporation rate ($\beta_{SA} C_{SA} / \sum \gamma$) of each cluster, calculated using DLPNO//DFT or DFT/6-31++G(d,p) at 278.15 K. [H_2SO_4] = 10^7 molecules cm^{-3} and [MBTCA] = 1 pptv.

themselves can drive new particle formation, under realistic atmospheric lower boundary layer conditions. Although effects such as anharmonicity[47] and hydration[48–52] can contribute to lowering the formation free energies of the clusters it is likely that the participation of other atmospheric vapours or other conditions is required. For instance Bianchi *et al.* showed that new particle formation *via* condensation of highly oxygenated molecules occurs in the free troposphere at a high-altitude research station Jungfraujoch (3580 m above sea level), in Switzerland.[53] Similarly, Kirkby *et al.* showed that ion-induced nucleation of highly oxidized molecules could occur without the presence of sulfuric acid.[54]

Previously, we studied sulfuric acid–pinic acid clusters, where it was also found that the molecular interaction was not strong enough to promote the formation of new particles at the lower boundary layer. The present study shows that the third carboxylic acid group in MBTCA yields more stable clusters than pinic acid can achieve, due to the potential of maximizing the amount of direct sulfuric acid–carboxylic acid interactions.

This suggests that in order for a terpene oxidation product to be able to participate in the initial steps in atmospheric new particle formation, several strong binding groups, such as carboxylic acid moieties, should be present, and the amount of sulfuric acid carboxylic acid interactions should be maximized to lead to the highest possible stabilization. Compounds which could fulfil these requirements could be yet unidentified exotic oxidation products of the first generation products pinonaldehyde, pinonic acid and pinic acid, or alternatively large dimers of monoterpene oxidation products.[55–57]

4 Conclusions

We have investigated the molecular interaction between the tricarboxylic acid MBTCA and sulfuric acid up to cluster sizes of $(MBTCA)_3(SA)_3$. The formation of the (SA)(MBTCA) heterodimer is found to be more thermodynamically favourable than either of the sulfuric acid or MBTCA homodimers. Adding MBTCA

molecules to the (SA)(MBTCA) complex is found more favourable than the corresponding addition of sulfuric acid molecules. The reaction free energies for forming $(MBTCA)_2(SA)_{2-3}$ clusters are found to be particularly favourable, with stability comparable to the sulfuric acid–dimethylamine complex. Generally, a high cluster stabilization is achieved when the amount of sulfuric acid–carboxylic acid interactions is maximized.

Guided by formation free energy surfaces and ambient concentrations we do, however, find that the MBTCA–SA clusters cannot at realistic vapour concentrations grow into larger stable clusters and will be susceptible to evaporation. Under atmospheric conditions it is thereby unlikely that MBTCA and sulfuric acid alone can drive the observed new particle formation events. Other stabilizing vapours are required. The $(MBTCA)_2(SA)_2$ and $(MBTCA)_2(SA)_3$ clusters are found to be most stable against evaporation, and if they can be formed, they could act as seeds for further growth by uptake of other stabilizing vapour molecules. This indicates that in order to reach a stable cluster, at least 8–12 direct sulfuric acid–carboxylic acid interactions are required.

Recently Mackenzie *et al.* showed that organic sulfuric anhydrides could be formed by cycloaddition of SO_3 to carboxylic acids.[58] Organic sulfuric anhydride formation would lower the vapor pressure of the compound compared to the corresponding carboxylic acid and significantly increase the interaction between clustering molecules. This could indicate that organic acids derived from α-pinene oxidation such as pinic acid and MBTCA could potentially get their ability to form molecular clusters significantly enhanced by forming an organic sulfuric anhydride.

Acknowledgements

J. E. thanks the Carlsberg foundation for financial support. We thank the Academy of Finland, Formas project 2015-749 and ERC project 257360-MOCAPAF, ERC 692891-DAMOCLES and ERC 278277-ATMOGAIN for funding. We thank the CSC-IT Center for Science in Espoo, Finland, for computational resources.

References

1 J. Almeida, S. Schobesberger, A. Kurten, I. K. Ortega, O. Kupiainen-Maatta, A. P. Praplan, A. Adamov, A. Amorim, F. Bianchi and M. Breitenlechner, *et al.*, *Nature*, 2013, **502**, 359–363.
2 R. Zhang, I. Suh, J. Zhao, D. Zhang, E. C. Fortner, X. Tie, L. T. Molina and M. J. Molina, *Science*, 2004, **304**, 1487–1490.
3 A. H. Goldstein and I. E. Galbally, *Environ. Sci. Technol.*, 2007, **41**, 1514–1521.
4 A. Guenther, C. N. Hewitt, D. Erickson, R. Fall, C. Geron, T. Graedel, P. Harley, L. Klinger, M. Lerdau, W. A. Mckay, T. Pierce, B. Scholes, R. Steinbrecher, R. Tallamraju, J. Taylor and P. Zimmerman, *J. Geophys. Res.*, 1995, **100**, 8873–8892.
5 J. H. Seinfeld and J. F. Pankow, *Annu. Rev. Phys. Chem.*, 2003, **54**, 121–140.
6 J. D. Crounse, L. B. Nielsen, S. Jørgensen, H. G. Kjaergaard and P. O. Wennberg, *J. Phys. Chem. Lett.*, 2013, **4**, 3513–3520.
7 T. Jokinen, M. Sipilä, S. Richters, V. Kerminen, P. Paasonen, F. Stratmann, D. Worsnop, M. Kulmala, M. Ehn, H. Herrmann and T. Berndt, *Angew. Chem., Int. Ed.*, 2014, **53**, 14596–14600.
8 T. Jokinen, T. Berndt, R. Makkonen, V.-M. Kerminen, H. Junninen, P. Paasonen, F. Stratmann, H. Herrmann, A. B. Guenther, D. R. Worsnop, M. Kulmala, M. Ehn and M. Sipilää, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 7123–7128.
9 T. Berndt, S. Richters, R. Kaethner, J. Voigtländer, F. Stratmann, M. Sipilää, M. Kulmala and H. Herrmann, *J. Phys. Chem. A*, 2015, **119**, 10336–10348.
10 M. P. Rissanen, T. Kurtén, M. Sipilä, J. A. Thornton, J. Kangasluoma, N. Sarnela, H. Junninen, S. Jørgensen, S. Schallhart and M. K. Kajos, *et al.*, *J. Am. Chem. Soc.*, 2014, **136**, 15596–15606.
11 M. P. Rissanen, T. Kurtén, M. Sipilä, J. A. Thornton, O. Kausiala, O. Garmash, H. G. Kjaergaard, T. Petäjä, D. R. Worsnop, M. Ehn and M. Kulmala, *J. Phys. Chem. A*, 2015, **119**, 4633–4650.
12 T. Kurtén, M. P. Rissanen, K. Mackeprang, J. A. Thornton, N. Hyttinen, S. Jørgensen and H. G. Kjaergaard, *J. Phys. Chem. A*, 2015, **119**, 11366–11375.
13 T. Hoffmann, R. Bandur, U. Marggraf and M. Linscheid, *J. Geophys. Res.*, 1998, **103**, 25569–25578.
14 T. S. Christoffersen, J. Hjorth, O. Horie, N. R. Jensen, D. Kotzias, L. L. Molander, P. Need, L. Ruppert, R. Winterhalter, A. Virkkula, K. Wirtz and B. R. Larsen, *Atmos. Environ.*, 1998, **32**, 1657–1661.
15 J. Yu, D. R. Cocker III, R. H. Griffin, R. C. Flagan and J. H. Seinfeld, *J. Atmos. Chem.*, 1999, **34**, 207–258.
16 M. Glasius, M. Duane and B. R. Larsen, *J. Chromatogr. A*, 1999, **833**, 121–135.
17 B. R. Larsen, D. D. Bella, M. Glacius, R. Winterhalter, N. R. Jensen and J. Hjorth, *J. Atmos. Chem.*, 2001, **38**, 231–275.
18 R. G. Prinn, R. F. Weiss, B. R. Millar, J. Haung, F. N. Alyea, D. M. Cunnold, P. J. Fraser, D. E. Hartley and P. G. Simmonds, *Science*, 1995, **269**, 187–192.
19 M. Hallquist, I. Wängberg and E. Ljungström, *Environ. Sci. Technol.*, 1997, **31**, 3166–3172.
20 M. Ehn, J. A. Thornton, E. Kleist, M. Sipila, H. Junninen, I. Pullinen, M. Springer, F. Rubach, R. Tillmann and B. Lee, *et al.*, *Nature*, 2014, **506**, 476–479.
21 F. Riccobono, S. Schobesberger, C. E. Scott, J. Dommen, I. K. Ortega, L. Rondo, J. Almeida, A. Amorim, F. Bianchi and M. Breitenlechner, *et al.*, *Science*, 2014, **344**, 717–721.
22 S. Schobesberger, H. Junninen, F. Bianchi, G. Lönn, M. Ehn, K. Lehtipalo, J. Dommen, S. Ehrhart, I. K. Ortega and A. Franchin, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 17223–17228.
23 J. Elm, N. Myllys, N. Hyttinen and T. Kurtén, *J. Phys. Chem. A*, 2015, **119**, 8414–8421.
24 J. Elm, N. Myllys, J. Luy, T. Kurtén and H. Vehkamäki, *J. Phys. Chem. A*, 2016, 2240–2249.

25 T. Kurtén, K. Tiusanen, P. Roldin, M. P. Rissanen, J. Luy, M. Boy, M. Ehn and N. M. Donahue, *J. Phys. Chem. A*, 2016, **120**, 2569–2582.
26 J. Elm, T. Kurtén, M. Bilde and K. V. Mikkelsen, *J. Phys. Chem. A*, 2014, **118**, 7892–7900.
27 J. W. Depalma, J. Wang, A. S. Wexler and M. V. Johnston, *J. Phys. Chem. A*, 2015, **119**, 11191–11198.
28 R. Szmigielski, J. D. Surratt, Y. Gómez-González, P. Van der Veken, I. Kourtchev, R. Vermeylen, F. Blockhuys, M. Jaoui, T. E. Kleindienst and M. Lewandowski, *et al.*, *Geophys. Res. Lett.*, 2007, **34**, L24811.
29 L. Müller, M. Reinnig, K. H. Naumann, H. Saathoff, T. F. Mentel, N. M. Donahue and T. Hoffmann, *Atmos. Chem. Phys.*, 2012, **12**, 1483–1496.
30 I. K. Ortega, N. M. Donahue, T. Kurtén, M. Kulmala, C. Focsa and H. Vehkamäki, *J. Phys. Chem. A*, 2016, **120**, 1452–1458.
31 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci and G. A. Petersson, *et al.*, *Gaussian 09, Revision B.01*, Gaussian, Inc., Wallingford CT, 2010.
32 J. Elm, M. Bilde and K. V. Mikkelsen, *J. Chem. Theory Comput.*, 2012, **8**, 2071–2077.
33 J. Elm, M. Bilde and K. V. Mikkelsen, *Phys. Chem. Chem. Phys.*, 2013, **15**, 16442–16445.
34 H. R. Leverentz, J. I. Siepmann, D. G. Truhlar, V. Loukonen and H. Vehkamäki, *J. Phys. Chem. A*, 2013, **117**, 3819–3825.
35 N. Bork, L. Du and H. G. Kjaergaard, *J. Phys. Chem. A*, 2014, **118**, 1384–1389.
36 C. Riplinger and F. Neese, *J. Chem. Phys.*, 2013, **138**, 034106.
37 C. Riplinger and F. Neese, *J. Chem. Phys.*, 2013, **139**, 134101.
38 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 73–78, DOI: 10.1002/wcms.81.
39 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, Avogadro: An advanced semantic chemical editor, visualization, and analysis platform, *J. Cheminf.*, 2012, **4**, 17.
40 J. Elm, M. Bilde and K. V. Mikkelsen, *J. Phys. Chem. A*, 2013, **117**, 6695–6701.
41 J. Elm, M. Fard, M. Bilde and K. V. Mikkelsen, *J. Phys. Chem. A*, 2013, **117**, 12990–12997.
42 J. Elm and K. V. Mikkelsen, *Chem. Phys. Lett.*, 2014, **615**, 26–29.
43 N. Myllys, J. Elm and T. Kurtén, *Comput. Theor. Chem.*, 2016, **1098**, 1–12.
44 N. Myllys, J. Elm, R. Halonen, T. Kurtén and H. Vehkamäki, *J. Phys. Chem. A*, 2016, **120**, 621–630.
45 R. Zhang, A. Khalizov, L. Wang, M. Hu and W. Xu, *Chem. Rev.*, 2012, **112**, 1957–2011.
46 D. R. Hanson and E. R. Lovejoy, *J. Phys. Chem. A*, 2006, **110**, 9525–9528.
47 B. Temelso, K. A. Archer and G. C. Shields, *J. Phys. Chem. A*, 2011, **115**, 12034–12046.
48 B. Temelso, T. N. Phan and G. C. Shields, *J. Phys. Chem. A*, 2012, **116**, 9745–9758.
49 D. E. Husar, B. Temelso, A. L. Ashworth and G. C. Shields, *J. Phys. Chem. A*, 2012, **116**, 5151–5163.
50 D. J. Bustos, B. Temelso and G. C. Shields, *J. Phys. Chem. A*, 2014, **118**, 7430–7441.
51 B. Temelso, T. E. Morrell, R. M. Shields, M. A. Allodi, E. K. Wood, K. N. Kirschner, T. C. Castonguay, K. A. Archer and G. C. Shields, *J. Phys. Chem. A*, 2012, **116**, 2209–2224.
52 H. Henschel, J. C. Acosta Navarro, T. Yli-Juuti, O. K. T. Olenius, I. K. Ortega, S. L. Clegg, T. Kurtén, I. Riipinen and H. Vehkamäki, *J. Phys. Chem. A*, 2014, **118**, 2599–2611.
53 F. Bianchi, J. Tröstl, H. Junninen, C. Frege, S. Henne, C. Hoyle, U. Molteni, E. Herrmann, A. Adamov, N. Bukowiecki, X. Chen, J. Duplissy, M. Gysel, M. Hutterli, J. Kangasluoma, J. Kontkanen, A. Kürten, H. E. Manninen, S. Münch, O. Peräkylä, T. Petäjä, L. Rondo, C. Williamson, E. Weingartner, J. Curtius, D. R. Worsnop, M. Kulmala, J. Dommen and U. Baltensperger, *Science*, 2016, **352**, 1109–1112.
54 J. Kirkby, J. Duplissy, K. Sengupta, C. Frege, H. Gordon, C. Williamson, M. Heinritzi, M. Simon, C. Yan, J. Almeida, J. Tröstl, T. Nieminen, I. K. Ortega, R. Wagner, A. Adamov, A. Amorim, A.-K. Bernhammer, F. Bianchi, M. Breitenlechner, S. Brilke, X. Chen, J. Craven, A. Dias, S. Ehrhart, R. C. Flagan, A. Franchin, C. Fuchs, R. Guida, J. Hakala, C. R. Hoyle, T. Jokinen, H. Junninen, J. Kangasluoma, J. Kim, M. Krapf, A. Kürten, A. Laaksonen, K. Lehtipalo, V. Makhmutov, S. Mathot, U. Molteni, A. Onnela, O. Peräkylä, F. Piel, T. Petäjä, A. P. Praplan, K. Pringle, A. Rap, N. A. D. Richards, I. Riipinen, M. P. Rissanen, L. Rondo, N. Sarnela, S. Schobesberger, C. E. Scott, J. H. Seinfeld, M. Sipilä, G. Steiner, Y. Stozhkov, F. Stratmann, A. Tomé, A. Virtanen, A. L. Vogel, A. C. Wagner, P. E. Wagner, E. Weingartner, D. Wimmer, P. M. Winkler, P. Ye, X. Zhang, A. Hansel, J. Dommen, N. M. Donahue, D. R. Worsnop, U. Baltensperger, M. Kulmala, K. S. Carslaw and J. Curtius, *Nature*, 2016, **533**, 521–526.
55 K. Kristensen, K. L. Enggrob, S. M. King, D. R. Worton, S. M. Platt, R. Mortensen, T. Rosenoern, J. D. Surratt, M. Bilde, A. H. Goldstein and M. Glasius, *Atmos. Chem. Phys.*, 2013, **13**, 3763–3776.
56 K. Kristensen, T. Cui, H. Zhang, A. Gold, M. Glasius and J. D. Surratt, *Atmos. Chem. Phys.*, 2014, **14**, 4201–4218.
57 X. Zhang, R. C. McVay, D. D. Huang, N. F. Dalleska, B. Aumont, R. C. Flagan and J. H. Seinfeld, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 14168–14173.
58 R. B. Mackenzie, C. T. Dewberry and K. R. Leopold, *Science*, 2015, **349**, 58–61.

