

Ion-induced nucleation of pure biogenic particles

Jasper Kirkby^{1,2}, Jonathan Duplissy^{3,4}, Kamalika Sengupta⁵, Carla Frege⁶, Hamish Gordon², Christina Williamson^{1,†}, Martin Heinritzi^{1,7}, Mario Simon¹, Chao Yan³, João Almeida^{1,2}, Jasmin Tröstl⁶, Tuomo Nieminen^{3,4}, Ismael K. Ortega⁸, Robert Wagner³, Alexey Adamov³, Antonio Amorim⁹, Anne-Kathrin Bernhammer^{7,10}, Federico Bianchi^{6,11}, Martin Breitenlechner^{7,10}, Sophia Brilke¹, Xuemeng Chen³, Jill Craven¹², Antonio Dias², Sebastian Ehrhart^{1,2}, Richard C. Flagan¹², Alessandro Franchin³, Claudia Fuchs⁶, Roberto Guida², Jani Hakala³, Christopher R. Hoyle^{6,13}, Tuija Jokinen³, Heikki Junninen³, Juha Kangasluoma³, Jaeseok Kim^{14,†}, Manuel Krapf⁶, Andreas Kürten¹, Ari Laaksonen^{14,15}, Katrianne Lehtipalo^{3,6}, Vladimir Makhmutov¹⁶, Serge Mathot², Ugo Molteni⁶, Antti Onnela², Otsu Peräkylä³, Felix Piel¹, Tuukka Petäjä³, Arnaud P. Praplan³, Kirsty Pringle⁵, Alexandru Rap⁵, Nigel A. D. Richards^{5,17}, Ilona Riipinen¹⁸, Matti P. Rissanen³, Linda Rondo¹, Nina Sarnela³, Siegfried Schobesberger^{3,†}, Catherine E. Scott⁵, John H. Seinfeld¹², Mikko Sipilä^{3,4}, Gerhard Steiner^{3,7,19}, Yuri Stozhkov¹⁶, Frank Stratmann²⁰, Antonio Tomé²¹, Annele Virtanen¹⁴, Alexander L. Vogel², Andrea C. Wagner¹, Paul E. Wagner¹⁹, Ernest Weingartner⁶, Daniela Wimmer^{1,3}, Paul M. Winkler¹⁹, Penglin Ye²², Xuan Zhang¹², Armin Hansel^{7,10}, Josef Dommen⁶, Neil M. Donahue²², Douglas R. Worsnop^{3,14,23}, Urs Baltensperger⁶, Markku Kulmala^{3,4}, Kenneth S. Carslaw⁵ & Joachim Curtius¹

Atmospheric aerosols and their effect on clouds are thought to be important for anthropogenic radiative forcing of the climate, yet remain poorly understood¹. Globally, around half of cloud condensation nuclei originate from nucleation of atmospheric vapours². It is thought that sulfuric acid is essential to initiate most particle formation in the atmosphere^{3,4}, and that ions have a relatively minor role⁵. Some laboratory studies, however, have reported organic particle formation without the intentional addition of sulfuric acid, although contamination could not be excluded^{6,7}. Here we present evidence for the formation of aerosol particles from highly oxidized biogenic vapours in the absence of sulfuric acid in a large chamber under atmospheric conditions. The highly oxygenated molecules (HOMs) are produced by ozonolysis of α -pinene. We find that ions from Galactic cosmic rays increase the nucleation rate by one to two orders of magnitude compared with neutral nucleation. Our experimental findings are supported by quantum chemical calculations of the cluster binding energies of representative HOMs. Ion-induced nucleation of pure organic particles constitutes a potentially widespread source of aerosol particles in terrestrial environments with low sulfuric acid pollution.

It is thought that aerosol particles rarely form in the atmosphere without sulfuric acid^{3,4}, except in certain coastal regions where iodine oxides are involved⁸. Furthermore, ions are thought to be relatively unimportant in the continental boundary layer, accounting for only around 10% of particle formation⁵. Sulfuric acid derives from anthropogenic and volcanic sulfur dioxide emissions as well as dimethyl sulfide from marine biota. However, typical daytime sulfuric acid concentrations (10^5 – 10^7 cm⁻³, or 0.004–0.4 parts per trillion by volume (p.p.t.v.) at standard conditions) are too low for sulfuric acid and water alone to account for the particle formation rates observed in the lower atmosphere⁹, so additional vapours are required to stabilize any embryonic sulfuric acid clusters against evaporation. Base species such as amines can do this and can explain part of atmospheric particle

nucleation¹⁰. It is well established that oxidation products of volatile organic compounds (VOCs) are important for particle growth¹¹, but whether their role in the smallest particles is in nucleation or growth alone has remained ambiguous^{4,12,13}. Recently, however, it has been shown that oxidized organic compounds do indeed help to stabilize sulfuric acid clusters and probably play a major role in atmospheric particle nucleation^{6,14,15}. We refer to these compounds as HOMs (highly oxygenated molecules) rather than ELVOCs (extremely low-volatility organic compounds)¹⁶ because the measured compounds span a wide range of low volatilities.

Here we report atmospheric particle formation solely from biogenic vapours. The data were obtained at the CERN CLOUD chamber (Cosmics Leaving Outdoor Droplets; see Methods for experimental details) between October 2012 and November 2013. In contrast with other works that have reported organic particle formation without intentional addition of sulfuric acid^{6,7}, here we measure the cluster chemistry and the role of ions, and rule out contamination.

Precursor VOCs in the atmosphere arise predominantly from natural sources such as vegetation and largely comprise isoprene (C₅H₈), monoterpenes (C₁₀H₁₆), sesquiterpenes (C₁₅H₂₄) and diterpenes (C₂₀H₃₂). Here we have studied α -pinene (C₁₀H₁₆) because it is the most abundant monoterpene, often exceeding 50 p.p.t.v. in the continental boundary layer¹⁷. We oxidized α -pinene by exposure to ozone and also to hydroxyl radicals (OH·) produced from ozone photolysis and secondary reactions. To measure the relative importance of these oxidants we also performed a few pure ozonolysis experiments (in which we removed OH· with a 0.1% H₂ scavenger) and a few pure hydroxyl experiments (in which we generated OH· by photolysis of gas-phase nitrous acid, HONO). Two nitrate chemical ionization atmospheric pressure interface time-of-flight (CI-APi-TOF) mass spectrometers measured neutral gas-phase compounds in the chamber (H₂SO₄ and HOMs). Therefore, for this study, HOMs are implicitly defined as oxidized organic compounds that can be detected by a nitrate CI-APi-TOF; related molecules with a lower oxidation state

¹Goethe University Frankfurt, Institute for Atmospheric and Environmental Sciences, 60438 Frankfurt am Main, Germany. ²CERN, CH-1211 Geneva, Switzerland. ³Department of Physics, University of Helsinki, FI-00014 Helsinki, Finland. ⁴Helsinki Institute of Physics, University of Helsinki, FI-00014 Helsinki, Finland. ⁵School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK. ⁶Paul Scherrer Institute, Laboratory of Atmospheric Chemistry, CH-5232 Villigen, Switzerland. ⁷Institute for Ion and Applied Physics, University of Innsbruck, 6020 Innsbruck, Austria. ⁸Onera—The French Aerospace Lab, F-91123 Palaiseau, France. ⁹SIM, University of Lisbon, 1849-016 Lisbon, Portugal. ¹⁰Ionicon Analytik GmbH, 6020 Innsbruck, Austria. ¹¹Institute for Atmospheric and Climate Science, ETH Zurich, CH-8092 Zurich, Switzerland. ¹²Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA. ¹³WSL Institute for Snow and Avalanche Research SLF, CH-7260 Davos, Switzerland. ¹⁴University of Eastern Finland, FI-70211 Kuopio, Finland. ¹⁵Finnish Meteorological Institute, FI-00010 Helsinki, Finland. ¹⁶Solar and Cosmic Ray Research Laboratory, Lebedev Physical Institute, 119991 Moscow, Russia. ¹⁷University of Leeds, National Centre for Earth Observation, Leeds LS2 9JT, UK. ¹⁸Department of Applied Environmental Science, University of Stockholm, SE-10961 Stockholm, Sweden. ¹⁹Faculty of Physics, University of Vienna, 1090 Vienna, Austria. ²⁰Leibniz Institute for Tropospheric Research, 04318 Leipzig, Germany. ²¹University of Beira Interior, 6201-001 Covilhã, Portugal. ²²Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA. ²³Aerodyne Research Inc., Billerica, Massachusetts 01821, USA. †Present addresses: CIRES, University of Colorado Boulder, Boulder, Colorado 80309, USA (C.W.); Arctic Research Center, Korea Polar Research Institute, Incheon 406-840, South Korea (J. Kim); Department of Atmospheric Sciences, University of Washington, Seattle, Washington 98195, USA (S.S.).

or different functional groups could be present in the chamber, but undetected by our nitrate chemical ionization set-up.

Before starting measurements, we carefully cleaned the CLOUD chamber (see Methods) and established extremely low contaminant concentrations: at 38% relative humidity and 278 K, the contaminants were below the detection limit for SO_2 (<15 p.p.t.v.) and H_2SO_4 ($<5 \times 10^4 \text{ cm}^{-3}$), and total organics (largely comprising high volatility C_1 – C_3 compounds) were below 150 p.p.t.v. Contaminants with a high proton affinity or a high gas-phase acidity can be detected as ions by the API-TOF operating in positive or negative mode, respectively, even at neutral molecule concentrations as low as 10^4 cm^{-3} . The API-TOF measured contaminant $\text{C}_5\text{H}_5\text{NH}^+$ (protonated pyridine) and contaminant NO_3^- to be the dominant positive and negative ions, respectively, before we added any trace gases to the chamber other than water vapour and ozone (Extended Data Fig. 1a, b). Despite its higher gas-phase acidity, we detected contaminant HSO_4^- at only 1% of the NO_3^- signal (Extended Data Fig. 1b), ruling out any contribution of sulfuric acid to the nucleation measurements. From previous studies and molecular analysis of the charged clusters (see below), the most abundant positive ion is likely to be contaminant ammonium (NH_4^+), but its mass is below the acceptance cut-off of the API-TOF as operated in this study.

Within a few minutes of the initial exposure of α -pinene to O_3 in the chamber, we detected gas-phase HOM monomers and dimers (Fig. 1a). Particles appeared shortly afterwards (Fig. 1b). HOM monomers (denoted E_1) broadly comprise highly oxidized $\text{C}_{8-10}\text{H}_{14,16}\text{O}_{6-12}$ species with an oxygen-to-carbon ratio (O/C) above about 0.6. HOM dimers (E_2) are two covalently bound monomers (see below), which generally have lower oxygen-to-carbon ratios, but, almost certainly, a lower volatility. For the present study we define E_1 (E_2) to be the summed HOM peaks in the mass/charge range $m/z = 235$ – 424 Th (425 – 625 Th), where $1 \text{ Th} = 1 \text{ Da}/e$ and e is the elementary charge. This definition excludes peaks in the E_1 mass band distinguished by an odd H number ($\text{C}_{10}\text{H}_{15}\text{O}_{6,8,10,12}$), which we assign to the $\text{RO}_2^\cdot$ peroxy radical. These m/z values include a contribution of 62 Th due to the NO_3^- ion from the CI-API-TOF ionizer. We define the total HOMs as the sum $\text{RO}_2^\cdot + \text{E}_1 + \text{E}_2$.

We measure high HOM molar yields (Extended Data Fig. 2): approximately 1.2% per hydroxyl radical ($\text{OH}^\cdot$) reaction with α -pinene, 3.2% per ozone reaction with α -pinene, and 2.9% from pure ozonolysis. We find a high E_2 yield from ozonolysis (10%–20% of total HOMs), but negligible E_2 yield from hydroxyl-initiated oxidation. Neutral trimers are close to the detection limit of the CI-API-TOF (below 0.1% of total HOMs). High yields of these same HOMs have previously been reported^{6,16}, although our ozonolysis yields are less than half those of ref. 16. For our experiments, α -pinene was in the range 0.1–2 parts per billion by volume (p.p.b.v.), with 20–40 p.p.b.v. of O_3 . The $\text{OH}^\cdot$ concentrations were $(0.5\text{--}0.8) \times 10^6 \text{ cm}^{-3}$ during ozonolysis experiments, and $(0.4\text{--}2) \times 10^5 \text{ cm}^{-3}$ during pure hydroxyl experiments with 0.5–3 p.p.b.v. of HONO .

This remarkably fast production of HOMs is likely to proceed via an autooxidation mechanism involving peroxy radicals^{16,18–20} (Extended Data Fig. 3). There is simply insufficient time for oxidation to proceed in multiple steps through stable intermediate molecules. Here, initial ozonolysis of an α -pinene molecule proceeds via a Criegee intermediate and further steps to form an $\text{RO}_2^\cdot$ radical, followed by several repeated cycles of intramolecular H abstraction and O_2 addition to re-form a new $\text{RO}_2^\cdot$ radical. We measure an $\text{RO}_2^\cdot$ fraction of total HOMs between 15% and 1% for HOMs from 0.1 p.p.t.v. to 10 p.p.t.v., respectively. A combination reaction of differently oxidized peroxy radicals explains the rapid high yield of covalently bound E_2 . The negligible E_2 yield from hydroxyl-initiated oxidation could result from additional NO_x chemistry that terminates the peroxy radicals before they can combine. Our theoretical calculations further indicate that E_2 must be covalently bound because the neutral molecular cluster formed from two monomers (denoted E_1E_1) is expected to be unstable (see below).

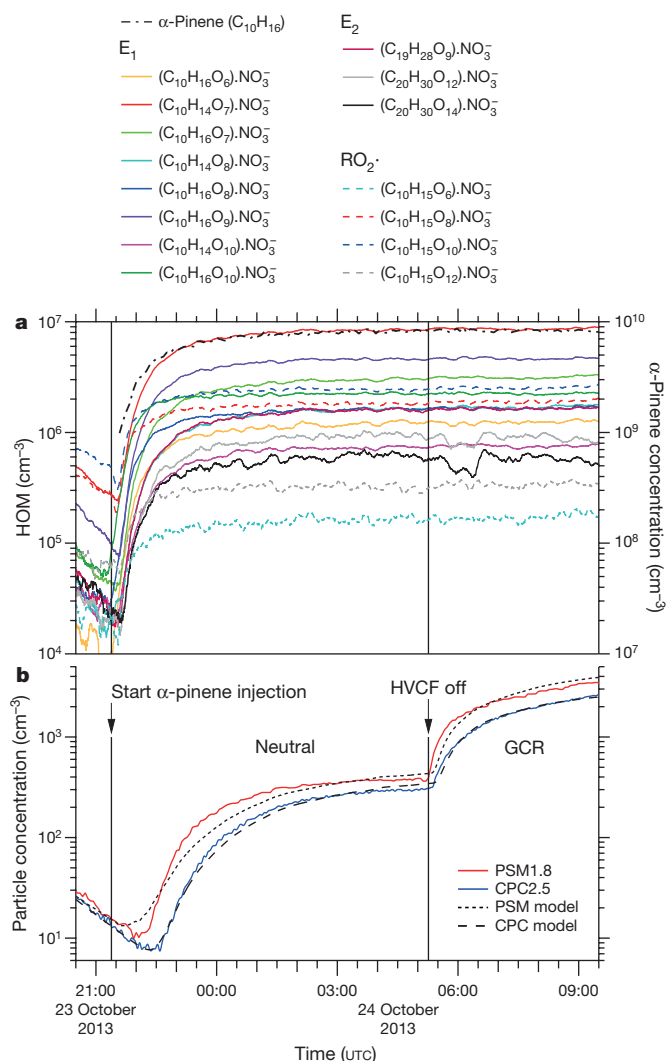


Figure 1 | Evolution of HOMs and particles during a typical run.

a, Evolution of selected HOM monomers (E_1), dimers (E_2) and peroxy radicals ($\text{RO}_2^\cdot$) at 300 p.p.t.v. α -pinene, 33 p.p.b.v. O_3 , zero H_2 or HONO , 38% relative humidity, 278 K and $[\text{H}_2\text{SO}_4] < 5 \times 10^4 \text{ cm}^{-3}$ (the same run as shown in Extended Data Fig. 4). The HOMs start to appear soon after the first injection of α -pinene into the chamber at 21:22, 23 October 2013. A HOM monomer is a highly oxygenated molecule derived from α -pinene ($\text{C}_{10}\text{H}_{16}$), and a HOM dimer is a covalently bound pair of monomers. Peroxy radicals are identified by an odd H number. The HOMs are charged with an NO_3^- ion in the CI-API-TOF mass spectrometer. The systematic scale uncertainty on the HOM concentrations is $+80\%/ -45\%$. **b**, Evolution of the particle number concentrations measured in the PSM1.8 (red curve) and CPC2.5 (blue curve) particle counters. The high-voltage clearing field (HVCF) was switched off at 05:16, 24 October 2013, marking the transition from neutral (ion-free) to GCR conditions in the chamber. A sharp increase in the rate of particle formation is seen, due to ion-induced nucleation of pure biogenic particles. However, no change occurs in the HOM concentrations (**a**), because these are predominantly neutral gas-phase molecules. The dotted and dashed curves in **b** show the PSM1.8 and CPC2.5 distributions, respectively, simulated for this run with the AEROCLOUD kinetic model, which is used to derive the experimental nucleation rates (see Methods).

We measured nucleation rates under neutral (J_n), Galactic cosmic ray (GCR; J_{gcr}) and π^+ beam (J_π) conditions, corresponding to ion-pair concentrations of around 0 cm^{-3} , 700 cm^{-3} and $3,000 \text{ cm}^{-3}$, respectively. This range spans atmospheric ion concentrations between ground level and 15-km altitude. The nucleation rate J_n describes the neutral rate alone, whereas J_{gcr} and J_π describe the sum of the neutral and ion-induced rates, $J_n + J_{\text{ion}}$. We determine the nucleation rates at

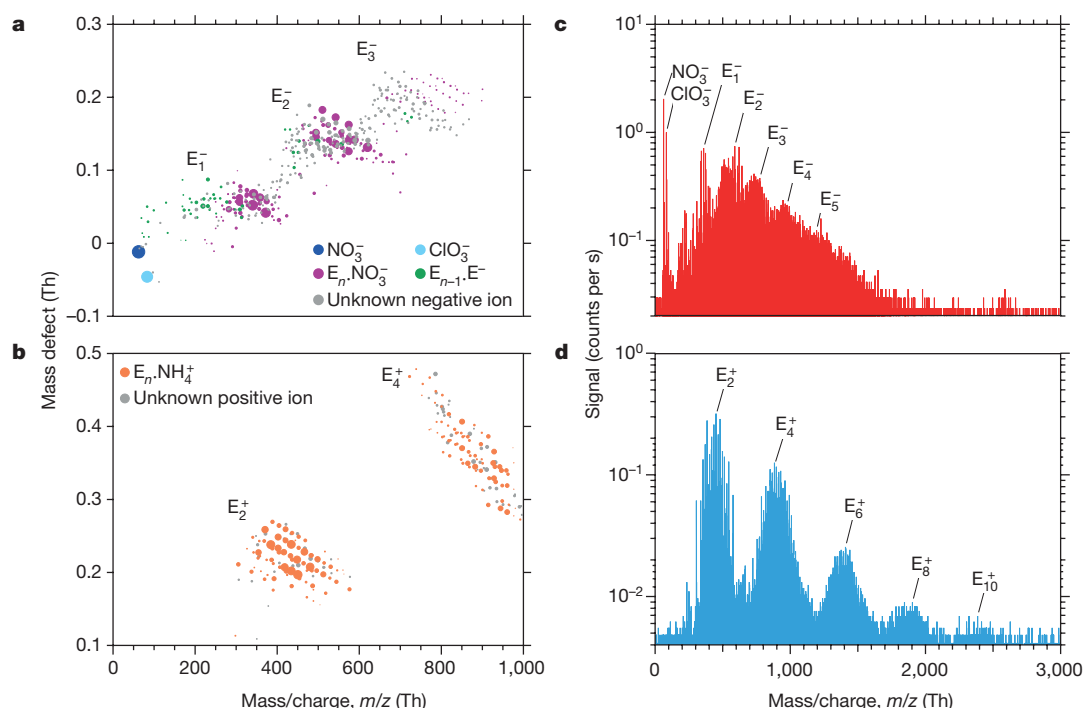


Figure 2 | Molecular composition and mass spectra of charged clusters during GCR nucleation events without sulfuric acid. **a, b,** Cluster mass defect (difference from integer mass) versus m/z of negatively (**a**) and positively (**b**) charged clusters measured with the API-TOF at 240 p.p.t.v. α -pinene, 34 p.p.b.v. O_3 , zero H_2 or HONO, 38% relative humidity, 278 K and $[H_2SO_4]$ below the detection limit ($5 \times 10^4 \text{ cm}^{-3}$). The values of J_{GCR} and total HOMs concentration are, respectively, $3.4 \text{ cm}^{-3} \text{ s}^{-1}$ and $1.7 \times 10^7 \text{ cm}^{-3}$ (**a**), and $3.3 \text{ cm}^{-3} \text{ s}^{-1}$ and $2.4 \times 10^7 \text{ cm}^{-3}$ (**b**). The mass bands are labelled according to the number of HOM monomer units in the cluster, E_n . Each circle represents a distinct molecular composition and its area represents the counts per second. The most highly oxidized compounds are located at the lower right-hand edge of each band.

1.7-nm mobility diameter, at which size a particle is generally considered to be stable against evaporation. To determine the nucleation rates, we fit the time-dependent particle concentrations with a numerical model that treats particle nucleation and growth kinetically at the molecular level (an example is shown in Fig. 1b; see Methods for further details).

A typical run sequence (Extended Data Fig. 4) begins by establishing ion-free conditions with a high-voltage clearing field and introducing α -pinene to the chamber, where it mixes with ozone. Particles then start to form and, after measuring J_n at steady-state α -pinene concentration, we turn off the high voltage and measure J_{GCR} under otherwise identical chamber conditions. A sharp enhancement of particle formation is seen when the high voltage was turned off (Extended Data Fig. 4b, e), due to ion-induced nucleation of both charge signs (Extended Data Figs 4c, d and 5).

Figure 2 shows the molecular composition and mass spectra of negatively and positively charged ions, monomers, dimers and clusters during ion-induced nucleation events. The dominant core ions in the clusters are identified as NH_4^+ , NO_3^- and E^- . Here E^- is inferred for negatively charged ions or clusters that contain only C, H and O; the E^- ion corresponds to a HOM of high gas-phase acidity. In contrast to negative clusters, the positive clusters nucleate only with dimers, producing distinct mass bands that are detected up to E_{10} in the API-TOF (Fig. 2c, d). This indicates the importance of dimers for pure biogenic nucleation. Dimers are expected to be less volatile than monomers, owing primarily to higher molecular weight, but also to additional functional groups. Our previously described definition for neutral gas-phase HOMs encompasses compounds with a wide range

The dark blue circle represents NO_3^- ions; the light blue circle represents ClO_3^- ions. Clusters with fully identified molecular composition are coloured according to their core ion: purple (NO_3^-), green (E^-) or orange (NH_4^+). Grey circles are unidentified clusters. **c, d,** Mass spectra from the same events for negative (**c**) and positive (**d**) clusters up to $m/z = 3,000$ Th. A particle of 1.7-nm mobility diameter has a mass of about 1,200 Th. The 'Nessie' plot (**d**) shows that positive-ion-induced nucleation involves HOM dimers alone ($E_1NH_4^+$ clusters are not seen owing to instrument tuning). The decreasing signal amplitude at larger masses is due to the lower concentration and decreasing detection efficiency of the API-TOF mass spectrometer (the efficiency versus m/z depends on the instrument tune and polarity).

of low volatilities^{19,21}, of which only a subset drive nucleation (ELVOCs, which comprise about 36% of measured total HOMs²¹). From the strong ion enhancement of nucleation we conclude that the API-TOF mass peaks above the dimer in Fig. 2 are clusters of ELVOC monomers and dimers. Although we can precisely determine their molecular composition ($C_xH_yO_z$), we can only infer their specific structure and functional groups.

We show the experimental neutral and GCR nucleation rates in Fig. 3 over the total HOMs range 0.1–10 p.p.t.v., which spans the range of atmospheric interest. Below 1 p.p.t.v. HOM, ionization at ground-level GCR intensities enhances the nucleation rate by between one and two orders of magnitude compared with neutral nucleation. At higher concentrations, the neutral and GCR nucleation rates converge because the ion-induced rate, J_{ion} , reaches the limit set by the GCR total ion production rate ($3.4 \text{ cm}^{-3} \text{ s}^{-1}$). Positive and negative clusters nucleate at comparable rates (an example is shown in Extended Data Fig. 5). Relative humidity has little effect on J_{GCR} over the range 6%–80% relative humidity, whereas J_n increases substantially at higher relative humidity (Extended Data Fig. 6).

The large GCR enhancement indicates that biogenic molecular clusters are relatively unstable unless an ion is present. A charged cluster is also likely to experience higher collision rates with HOMs because they are expected to have high electric polarizability and, depending on their structure, large dipole moments. We further investigated the dependence on ion species by adding small amounts of SO_2 to the chamber, up to around 1,000 p.p.t.v. When $[H_2SO_4]$ exceeds about $1 \times 10^5 \text{ cm}^{-3}$, the major negative ion species shift to HSO_4^- , SO_5^- and SO_4^- (Extended Data Fig. 1c), owing to their lower proton affinity (higher gas-phase

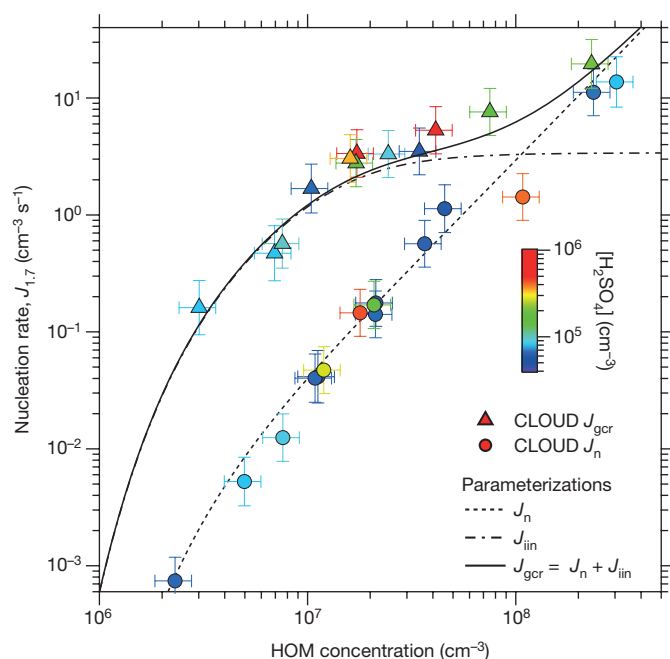


Figure 3 | Pure biogenic nucleation rates versus HOM concentration. Neutral (J_n ; circles) and GCR (J_{gr} ; triangles) nucleation rates versus total HOMs concentration ($\text{RO}_2 + \text{E}_1 + \text{E}_2$). The fraction of total HOMs that participate in nucleation (ELVOCs) is about 36% (ref. 21). The experimental conditions are 10–1,300 p.p.t.v. α -pinene (for measurements below $J_{1.7} = 10 \text{ cm}^{-3} \text{ s}^{-1}$), 30–35 p.p.b.v. O_3 , zero H_2 or HONO, 38% relative humidity, 278 K and $< 8 \times 10^5 \text{ cm}^{-3} \text{ H}_2\text{SO}_4$. The colour scale shows $[\text{H}_2\text{SO}_4]$; purple and blue points correspond to contaminant level (below the detection threshold); other colours correspond to measurements after SO_2 was added to the chamber. The fitted curves show parameterizations (described in Methods) for J_n (dashed), J_{gr} (solid) and ion-induced nucleation ($J_{iin} = J_{gr} - J_n$; dot-dashed). The J_{iin} parameterization assumes that the nucleation rate falls steeply at HOM concentrations below the experimental measurements, following a similar slope to that for J_n . The bars indicate 1σ total errors, although the overall systematic scale uncertainty of $+80\%/ -45\%$ on the HOM concentration is not shown.

acidity) than contaminant compounds. However, the nucleation rates with sulfur ion species remain unchanged (Fig. 3). Taken together, our observations therefore show that ubiquitous ion species can stabilize embryonic biogenic clusters. However, we do not observe chlorine in nucleating clusters, even though contaminant chlorine ion species are present (Fig. 2 and Extended Data Fig. 1), which indicates that not all ions have a suitable chemical structure to bond strongly with the oxidized organic compounds²².

Figure 4 shows the CLOUD biogenic nucleation rates extended to $[\text{H}_2\text{SO}_4] = 6 \times 10^6 \text{ cm}^{-3}$ and compared with atmospheric boundary-layer observations^{3,4,23,24}. Biogenic nucleation rates show no significant dependence on sulfuric acid concentration over this range (that is, within the experimental measurement errors, the nucleation rate is consistent with zero dependency on sulfuric acid concentration). This finding sharply contrasts with base-stabilized nucleation of sulfuric acid in the presence of ammonia⁹ or amines¹⁰, where nucleation rates at 1.7 nm show a steep dependency on $[\text{H}_2\text{SO}_4]$ above 10^6 cm^{-3} . Comparison of the atmospheric observations (Fig. 4) with our measurements therefore suggests that nucleation in the lower atmosphere may involve a mixture of two distinct mechanisms. The first, which is more important in polluted environments, involves nucleation of sulfuric acid and water together with a combination of amines or ammonia with oxidized organics, and has a strong dependence on sulfuric acid. The second, which is more important in pristine environments, involves nucleation of pure organic particles and depends on only oxidized organics and ions.

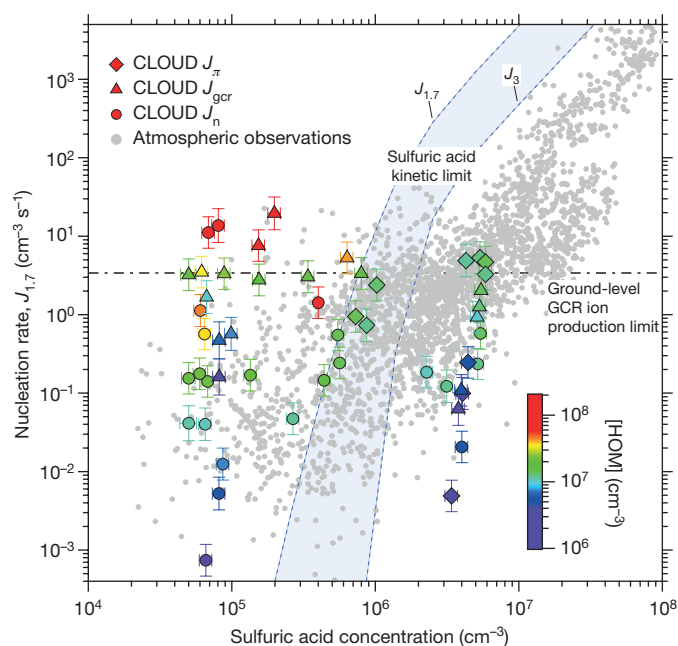


Figure 4 | Experimental and atmospheric nucleation rates versus H_2SO_4 concentration. CLOUD measurements of the neutral (J_n ; circles), GCR (J_{gr} ; triangles) and π beam (J_{pi} ; diamonds) biogenic nucleation rates at 1.7 nm ($J_{1.7}$) versus $[\text{H}_2\text{SO}_4]$. The CLOUD experimental conditions are 10–1,300 p.p.t.v. α -pinene (for measurements below $J_{1.7} = 10 \text{ cm}^{-3} \text{ s}^{-1}$), 25–35 p.p.b.v. O_3 , zero H_2 or HONO, 20%–40% relative humidity and 278 K. Measurements below $1 \times 10^5 \text{ cm}^{-3}$ for $[\text{H}_2\text{SO}_4]$ are near to the detection limit of the CI-API-TOF and should be considered as upper-estimates (to avoid overlap, some data points at the H_2SO_4 detection limit are displaced by up to $1 \times 10^4 \text{ cm}^{-3}$). The total HOMs concentration from α -pinene oxidation is indicated by the colour scale. Observations of particle formation in the atmospheric boundary layer (mainly at 3-nm threshold size) are indicated by small grey circles^{3,4,23,24}. Following convention, the H_2SO_4 concentration refers to monomers alone; that is, H_2SO_4 bound in molecular clusters is not included. The kinetic upper limit on sulfuric acid nucleation is indicated by the blue band, which is bounded by dashed lines indicating $J_{1.7}$ and J_3 . This band assumes the CLOUD condensation sink, which is comparable to that of a pristine atmosphere. The upper limit on J_{iin} from the GCR ion-pair production rate at ground level is indicated by the dot-dashed line. The bars indicate 1σ total errors, although the overall $+50\%/ -33\%$ systematic scale uncertainty on $[\text{H}_2\text{SO}_4]$ is not shown.

To gain further insight into the stability of initial neutral and charged clusters of highly oxidized biogenic molecules, we calculated their Gibbs free energies of formation, ΔG , using quantum chemical methods (see Methods). For this study we chose $\text{C}_{10}\text{H}_{14}\text{O}_7$ and $\text{C}_{20}\text{H}_{30}\text{O}_{14}$ as E_1 and E_2 surrogates, respectively (Extended Data Fig. 7). We observe these compounds both in the gas (Fig. 1) and particle phases in the CLOUD chamber. We show proposed formation mechanisms and structures^{19,20} in Extended Data Fig. 3. Our calculations, summarized in Extended Data Table 1 and Extended Data Fig. 8, confirm that ELVOC clusters formed with an E_1^- , HSO_4^- , NO_3^- or NH_4^+ ion are expected to be stable (that is, their growth rate exceeds the evaporation rate) at around 0.1 p.p.t.v. ELVOC, or below. In contrast, the initial neutral clusters are weakly bound and so neutral nucleation is expected to be weaker. Although limited to a single surrogate pair, our theoretical calculations thus provide independent support for the experimental measurements.

Comparisons with atmospheric observations should be considered as preliminary because our measurements were made at only one temperature, with a single monoterpene, in the absence of isoprene and mostly in the absence of NO_x , which can influence HOM yields. Nevertheless, our results may provide fresh insights into several seemingly disparate

phenomena associated with low atmospheric concentrations of sulfuric acid. First, pure HOM nucleation could provide a mechanism to account for nucleation-mode particles observed at night-time, under low- $[\text{H}_2\text{SO}_4]$ conditions^{25,26}. Second, although observations are rare, nucleation-mode particles are seen in the Amazon²⁷, where SO_2 levels are extremely low (20–30 p.p.t.v.). Peak particle concentrations often occur at sunrise and sunset²⁷, and appear to be associated with rain, which reduces the aerosol condensation sink and may generate high ion concentrations by evaporation of charged droplets at the Rayleigh limit. Third, pure biogenic nucleation could explain new particle formation observed in the upper troposphere in cloud outflows depleted of SO_2 , such as over the Amazon^{27–29}. Low-solubility biogenic precursor vapours can be efficiently convected inside clouds to high altitudes where HOMs will form in the cloud outflows on exposure to oxidants, and nucleation is likely to be enhanced by the low temperatures. Fourth, since high HOM yields are also found from other organic compounds with an endocyclic double bond such as cyclohexene¹⁶, pure HOM nucleation involving anthropogenic organic precursors could be expected when $[\text{H}_2\text{SO}_4]$ is low³⁰. Finally, ion-induced pure biogenic nucleation might shed new light on the long-standing question of a physical mechanism for solar-climate variability in the pristine pre-industrial climate^{31,32}.

Direct observational evidence of pure biogenic nucleation has not been reported so far, owing to atmospheric pollution or lack of suitable instrumentation. The pure biogenic mechanism is likely to dominate nucleation in pristine terrestrial regions such as tropical rainforests or at higher altitudes above forests in convective cloud outflows. Pure biogenic nucleation might also take place over forested areas at high northern latitudes during periods of especially low pollution. Identification of pure biogenic nucleation in the atmosphere will require simultaneous measurements with several newly developed mass spectrometers, API-TOF (for molecular composition of ions and nucleating charged clusters) and CI-API-TOF (gas-phase HOMs and H_2SO_4), together with standard instruments such as low-threshold particle counters, PTR-TOF (precursor organic vapours) and NAIS (size spectra of ions and charged particles).

In summary, we find that highly oxidized organic compounds play a role in atmospheric particle nucleation comparable to that of sulfuric acid; together with a suitable stabilizing agent, each has sufficiently low volatility to form new particles in the lower atmosphere at vapour concentrations near 10^7 cm^{-3} . The stabilizing agent for pure biogenic particles is a suitable ion, whereas for sulfuric acid particles the stabilizing agents are amines, or ammonia with oxidized organics. Ion-induced nucleation of pure biogenic particles may have important consequences for pristine climates because it provides a mechanism by which nature produces particles without pollution. This could raise the baseline aerosol state of the pristine pre-industrial atmosphere and so could reduce the estimated anthropogenic radiative forcing from increased aerosol-cloud albedo over the industrial period.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 July 2015; accepted 16 March 2016.

1. Boucher, O. *et al.* in *Climate Change 2013: The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. *et al.*) 571–658 (Cambridge Univ. Press, 2013).
2. Merikanto, J., Spracklen, D. V., Mann, G. W., Pickering, S. J. & Carslaw, K. S. Impact of nucleation on global CCN. *Atmos. Chem. Phys.* **9**, 8601–8616 (2009).
3. Kuang, C., McMurry, P. H., McCormick, A. V. & Eisele, F. L. Dependence of nucleation rates on sulfuric acid vapor concentration in diverse atmospheric locations. *J. Geophys. Res. Atmos.* **113**, D10209 (2008).
4. Kulmala, M. *et al.* Direct observations of atmospheric aerosol nucleation. *Science* **339**, 943–946 (2013).
5. Hirsikko, A. *et al.* Atmospheric ions and nucleation: a review of observations. *Atmos. Chem. Phys.* **11**, 767–798 (2011).

6. Zhao, J., Ortega, J., Chen, M., McMurry, P. H. & Smith, J. N. Dependence of particle nucleation and growth on high-molecular-weight gas-phase products during ozonolysis of α -pinene. *Atmos. Chem. Phys.* **13**, 7631–7644 (2013).
7. Gao, S. *et al.* Low-molecular-weight and oligomeric components in secondary organic aerosol from the ozonolysis of cycloalkenes and α -pinene. *J. Phys. Chem. A* **108**, 10147–10164 (2004).
8. O'Dowd, C. D. *et al.* Marine aerosol formation from biogenic iodine emissions. *Nature* **417**, 632–636 (2002).
9. Kirkby, J. *et al.* Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation. *Nature* **476**, 429–433 (2011).
10. Almeida, J. *et al.* Molecular understanding of sulphuric acid–amine particle nucleation in the atmosphere. *Nature* **502**, 359–363 (2013).
11. Riipinen, I. *et al.* Organic condensation: a vital link connecting aerosol formation to cloud condensation nuclei (CCN) concentrations. *Atmos. Chem. Phys.* **11**, 3865–3878 (2011).
12. Zhang, R. *et al.* Atmospheric new particle formation enhanced by organic acids. *Science* **304**, 1487–1490 (2004).
13. Metzger, A. *et al.* Evidence for the role of organics in aerosol particle formation under atmospheric conditions. *Proc. Natl Acad. Sci. USA* **107**, 6646–6651 (2010).
14. Schobesberger, S. *et al.* Molecular understanding of atmospheric particle formation from sulfuric acid and large oxidized organic molecules. *Proc. Natl Acad. Sci. USA* **110**, 17223–17228 (2013).
15. Riccobono, F. *et al.* Oxidation products of biogenic emissions contribute to nucleation of atmospheric particles. *Science* **344**, 717–721 (2014).
16. Ehn, M. *et al.* A large source of low-volatility secondary organic aerosol. *Nature* **506**, 476–479 (2014).
17. Guenther, A. B. *et al.* The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions. *Geosci. Model Dev.* **5**, 1471–1492 (2012).
18. Crounse, J. D., Nielsen, L. B., Jørgensen, S., Kjaergaard, H. G. & Wennberg, P. O. Autoxidation of organic compounds in the atmosphere. *J. Phys. Chem. Lett.* **4**, 3513–3520 (2013).
19. Zhang, X. *et al.* Formation and evolution of molecular products in α -pinene secondary organic aerosol. *Proc. Natl Acad. Sci. USA* **112**, 14168–14173 (2015).
20. Kurtén, T. *et al.* Computational study of hydrogen shifts and ring-opening mechanisms in α -pinene ozonolysis products. *J. Phys. Chem. A* **119**, 11366–11375 (2015).
21. Tröstl, J. *et al.* The role of low-volatility organic compounds in initial particle growth in the atmosphere. *Nature* **533**, <http://dx.doi.org/10.1038/nature18271> (2016).
22. Kathmann, S. M., Schenter, G. K. & Garrett, B. C. Ion-induced nucleation: the importance of chemistry. *Phys. Rev. Lett.* **94**, 116104 (2005).
23. Paasonen, P. *et al.* On the roles of sulphuric acid and low-volatility organic vapours in the initial steps of atmospheric new particle formation. *Atmos. Chem. Phys.* **10**, 11223–11242 (2010).
24. Sihto, S.-L. *et al.* Atmospheric sulphuric acid and aerosol formation: implications from atmospheric measurements for nucleation and early growth mechanisms. *Atmos. Chem. Phys.* **6**, 4079–4091 (2006).
25. Suni, T. *et al.* Formation and characteristics of ions and charged aerosol particles in a native Australian eucalyptus forest. *Atmos. Chem. Phys.* **8**, 129–139 (2008).
26. Lee, S.-H. *et al.* Observations of nighttime new particle formation in the troposphere. *J. Geophys. Res. Atmos.* **113**, D10210 (2008).
27. Martin, S. T. *et al.* Sources and properties of Amazonian aerosol particles. *Rev. Geophys.* **48**, RG2002 (2010).
28. Kulmala, M. *et al.* Deep convective clouds as aerosol production engines: role of insoluble organics. *J. Geophys. Res. Atmos.* **111**, D17202 (2006).
29. Ekman, A. M. L. *et al.* Do organics contribute to small particle formation in the Amazonian upper troposphere? *Geophys. Res. Lett.* **35**, L17810 (2008).
30. Bianchi, F. *et al.* New particle formation in the free troposphere: a question of chemistry and timing. *Science* **352**, <http://dx.doi.org/10.1126/science.aad5456> (2016).
31. Herschel, W. Observations tending to investigate the nature of the Sun, in order to find the causes or symptoms of its variable emission of light and heat; with remarks on the use that may possibly be drawn from solar observations. *Philos. Trans. R. Soc. Lond.* **91**, 265–318 (1801).
32. Kirkby, J. Cosmic rays and climate. *Surv. Geophys.* **28**, 333–375 (2007).

Acknowledgements We thank CERN for supporting CLOUD with important technical and financial resources, and for providing a particle beam from the CERN Proton Synchrotron. We also thank P. Carrie, L.-P. De Menezes, J. Dumollard, F. Josa, I. Krasin, R. Kristic, A. Laassiri, O. S. Maksumov, B. Marichy, H. Martinati, S. V. Mizin, R. Sitals, A. Wasem and M. Wilhelmsson for their contributions to the experiment. We thank the CSC Centre for Scientific Computing in Espoo, Finland for computer time. This research has received funding from the EC Seventh Framework Programme (Marie Curie Initial Training Network MC-ITN CLOUD-TRAIN no. 316662, EU Horizon 2020 Marie Curie grant no. 656994, ERC-Consolidator grant NANODYNAMITE no. 616075 and ERC-Advanced grant ATMNUCLE no. 227463), the German Federal Ministry of Education and Research (project no. 01LK1222A), the Swiss National Science Foundation (project nos 200020_135307, 200021_140663, 206021_144947/1 and 20FI20_149002/1), the Academy of Finland (Center of Excellence project no. 1118615), the Academy of Finland (135054, 133872,

251427, 139656, 139995, 137749, 141217, 141451), the Finnish Funding Agency for Technology and Innovation, the Väisälä Foundation, the Nesslering Foundation, the Austrian Science Fund (FWF; project no. L593), the Portuguese Foundation for Science and Technology (project no. CERN/FP/116387/2010), the Swedish Research Council, Vetenskapsrådet (grant 2011-5120), the Presidium of the Russian Academy of Sciences and Russian Foundation for Basic Research (grant 12-02-91522-CERN), the UK Natural Environment Research Council (grant NE/K015966/1), the Royal Society (Wolfson Merit Award), the US National Science Foundation (grants AGS1136479, AGS1447056 and CHE1012293), Caltech ESE Grant (Davidow Foundation), Dreyfus Award EP-11-117, the French National Research Agency (ANR), the Nord-Pas de Calais, and the European Funds for Regional Economic Development (FEDER, Labex-Cappa, ANR-11-LABX-0005-01).

Author Contributions J.A., H.G., A.K., T.N., J.T. and C.W. analysed the nucleation rates; C.Fr. analysed the API-TOF charged clusters; M.H., M.Sim. and C.Y. performed the CI-API-TOF HOM and H₂SO₄ analyses; A.-K.B. analysed the PTR-TOF α -pinene; J.H.S. and X.Z. analysed the ELVOC structures and formation mechanisms; I.K.O. performed the quantum chemical calculations; A.Ad., J.A., A.Am., A.-K.B., F.B., M.B., S.B., J.Cu., J.Cr., A.D., J.Do., J.Du., S.E., C.Fr., C.Fu., H.G., M.H., C.R.H., T.J., H.J., J.Ka., J. Kim, J.Kir., M.Kr., A.K., K.L., V.M., U.M., T.N., F.P., T.P., A.P.P., M.P.R., N.S., K.S., M.Sim., M.Sip., G.S., A.T., J.T., A.W., D.W., R.W., C.W.,

C.Y. and P.Y. collected the data and contributed to the analysis; K.S.C., H.G., K.P., A.R., N.A.D.R., K.S. and C.E.S. evaluated the atmospheric relevance; J.Kir. wrote the manuscript; J.A., J.Do., N.M.D., C.Fr., H.G., M.H., J.H.S., M.Sim., C.W., R.W., C.Y. and X.Z. contributed to Methods and Extended Data; and U.B., K.S.C., J.Cu., J.Do., N.M.D., R.C.F., A.H., J.Kir., M.Ku., J.H.S. and D.R.W. contributed to data interpretation and editing of manuscript. All authors contributed to the development of the CLOUD facility and analysis instruments, and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.Kir. (jasper.kirkby@cern.ch).



This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons licence, users will need to obtain permission from the licence holder to reproduce the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

METHODS

Overview of the CLOUD facility. The CLOUD experiment at CERN is designed to study the effects of cosmic rays on aerosols, cloud droplets and ice particles, under precisely controlled laboratory conditions. The 3-m-diameter stainless-steel CLOUD chamber and its gas system have been built to the highest technical standards of cleanliness and performance. The CLOUD chamber is periodically cleaned by rinsing the walls with ultra-pure water, followed by heating to 373 K and flushing at a high rate with humidified synthetic air and elevated ozone (several parts per million by volume). Contaminant levels of condensable vapours are in the sub-p.p.t.v. range. The high cleanliness of the chamber, together with its large volume (26.1 m³) and highly stable operating conditions, allows particle formation to be studied under atmospheric conditions at nucleation rates between about 0.001 cm⁻³ s⁻¹ and 100 cm⁻³ s⁻¹. The loss rate of condensable vapours and particles onto the chamber walls is comparable to the ambient condensation sink of the pristine boundary layer.

Ion production in the chamber can be controlled using an internal electric clearing field (which creates an ion-free environment), GCRs or an adjustable π^+ beam^{9,33} from the CERN Proton Synchrotron. The π^+ beam is de-focused to a transverse size of about 1.5 m × 1.5 m when it passes through the CLOUD chamber. With the electric field set to zero, the equilibrium ion-pair concentration in the chamber due to GCRs is around 700 cm⁻³. With the π^+ beam, this can be increased to any value up to about 3,000 cm⁻³. Hence, ion concentrations corresponding to any altitude in the troposphere can be generated in the CLOUD chamber.

The experiment has precise control of the trace vapours inside the chamber and also of the environmental temperature between 300 K and 203 K. Uniform mixing is achieved with magnetically coupled stainless-steel fans mounted at the top and bottom of the chamber. The characteristic gas mixing time in the chamber is a few minutes, depending on the fan speeds. Photochemical processes are initiated by illumination with an ultraviolet fibre-optic system, providing highly stable gas-phase reactions with a precise start time. The contents of the chamber are continuously analysed by instruments connected to sampling probes that project into the chamber. The sampling analysers are tailored for each experimental campaign, but typically comprise around 30–35 instruments, of which up to 10 are mass spectrometers.

Summary of analysing instruments. For the results reported here, the analysing instruments attached to the chamber included a chemical ionization mass spectrometer (CIMS) for H₂SO₄ concentration³⁴; an atmospheric pressure interface time-of-flight (API-TOF; Aerodyne Research Inc. and ToFwerk AG)³⁵ mass spectrometer for molecular composition of positively or negatively charged ions and clusters; two chemical ionization atmospheric pressure interface time-of-flight (CI-API-TOF; Aerodyne Research Inc. and ToFwerk AG)^{36,37} mass spectrometers for molecular composition and concentration of neutral gas-phase H₂SO₄ and HOMs; a proton transfer reaction time-of-flight (PTR-TOF; Ionicon Analytik GmbH)³⁸ mass spectrometer for organic vapours; a neutral cluster and air ion spectrometer (NAIS; Aired Ltd)³⁹ for concentrations of positive ions, negative ions and charged clusters in the range 1–40 nm; a nano-radial differential mobility analyser (nRDMA)⁴⁰ and a nano scanning mobility particle sizer (nano-SMPS) for particle size spectra; and several condensation particle counters (CPCs) with 50% detection efficiency thresholds between 1 nm and 4 nm: two Airmodus A09 particle size magnifiers, PSM⁴¹, (one fixed-threshold and the other scanning), two diethylcylene glycol CPCs, DEG-CPC^{42,43}, a butanol TSI 3776 CPC and a water TSI 3786 CPC (TSI Inc.).

Additional gas analysers included dew-point sensors (EdgeTech), sulfur dioxide (Thermo Fisher Scientific, Inc. 42i-TLE) and ozone (Thermo Environmental Instruments TEI 49C). For certain tests, HONO vapour was supplied to the chamber and photolysed with ultraviolet light to produce OH· in the absence of O₃. The gaseous HONO was generated by continual mixing of H₂SO₄ with NaNO₂ (ref. 44) in a specially designed stainless-steel reactor, and then steadily flowed into the chamber. The HONO analyser involved a specially designed probe that passed samples of air from the chamber through a solution of H₂SO₄ and sulfanilamide, which was then analysed online with a long path absorption photometer (LOPAP)⁴⁵.

Determination of the nucleation and growth rates. The nucleation rates (in cm⁻³ s⁻¹) were measured under neutral (J_n), ground-level GCR (J_{GCR}) and π^+ beam (J_π) conditions. Neutral nucleation rates are measured with the clearing field electrodes set to ±30 kV, which establishes an electric field of about 20 kV m⁻¹ in the chamber. This completely suppresses ion-induced nucleation because, under these conditions, small ions or molecular clusters are swept from the chamber in about 1 s. Because all of the nucleation and growth processes under consideration take place on substantially longer timescales, neutral nucleation rates can be measured with zero background from ion-induced nucleation. For GCR and π^+ beam conditions, the electric field was set to zero, leading to equilibrium ion-pair concentrations around 700 cm⁻³ and 3,000 cm⁻³, respectively. The nucleation rate

J_n measures the neutral rate alone, whereas J_{GCR} and J_π measure the sum of the neutral and ion-induced nucleation rates, $J_n + J_{\text{ion}}$.

The nucleation rates reported here were obtained primarily with the Airmodus scanning PSM at 1.8-nm threshold (PSM1.8) and the TSI 3776 CPC (CPC2.5), nominally 2.5-nm threshold, but measured at 3.2-nm threshold with WO_x particles⁴⁶. The nucleation rates $J_{1.7}$ are determined at 1.7-nm mobility diameter (1.4-nm mass diameter), at which size a particle is normally considered to be above its critical size and, therefore, thermodynamically stable. The critical size corresponds to the cluster size at which the evaporation and growth rates are equal. It varies with temperature, chemical species, charge and vapour concentrations, and may even be absent when evaporation rates are highly suppressed, such as for sulfuric acid–dimethylamine clusters^{10,37}. Our measurements indicate that the smallest neutral HOM clusters are relatively unstable; therefore, 1.7 nm, which is equivalent to around 5 HOM monomer units, is a reasonable size at which to derive the experimental nucleation rates.

AEROCLOUD model. To determine nucleation rates at 1.7 nm, the time-dependent particle concentrations measured with the PSM1.8 and CPC2.5 are fitted with a simplified numerical model (AEROCLOUD) that treats particle nucleation and growth kinetically at the molecular level. The model uses HOM monomer, HOM dimer and H₂SO₄ production rates derived from the CI-API-TOF experimental data. The measured HOM production rates are scaled by a factor of 1.8 to match the observed particle appearance times and growth rates. This scaling results in good agreement of the model with the experimental data over the full experimental range of HOM concentrations. The scaling factor is within the systematic measurement uncertainty of the CI-API-TOF, and could arise if a nitrate CI-API-TOF does not detect all the HOMs that contribute to particle growth.

Primary ions from GCRs are generated in the model at the known rate of $q = 1.7$ ion pairs per cubic centimetre per second. A fixed parameter of the model, f_c , accounts for the charge sign asymmetry due to differences in the diffusional loss rates of positive and negative primary ions to the chamber walls:

$$q_+ = f_c(2q) \\ q_- = (1 - f_c)(2q)$$

The parameter f_c is determined by the experimentally measured positive and negative ion concentrations in the NAIS to have the value 0.52.

Molecules and particles collide kinetically, and cluster with each other. The model uses a reduced clustering probability (termed a ‘sticking probability’ below) to account for unstable small clusters, rather than allowing clusters to evaporate once they have formed. This greatly increases the speed of the computation. If the particle formed by a collision exceeds a certain size (corresponding to around 1.7-nm mobility diameter for pure biogenic clusters; see below), then it is assumed to be effectively stable and subsequently grows at near the kinetic limit. The particle growth rate between the PSM1.8 and CPC2.5 is therefore implicitly treated in the model essentially as kinetically limited growth by particle coagulation plus HOM and H₂SO₄ vapour condensation. Particles grow through size bins that are linearly spaced for small sizes and logarithmically spaced from about 2 nm to a maximum size of 400 nm. The time-steps for clustering processes range from 0.9 s to 10 s, depending on the conditions of the experimental run under analysis. The time-step is 10 s for all other processes (for example, updates of gas concentrations, high-voltage clearing-field changes, fan changes, and particle losses due to dilution of the chamber contents or diffusion to the walls). The density of the pure HOM clusters is fixed at 1.3 g cm⁻³, and at 1.85 g cm⁻³ for a pure H₂SO₄ cluster.

For neutral–neutral collisions, the number of particles in size bins 1 and 2 that coagulate in a time interval Δt to produce a particle of mass m_{12} is:

$$n_{12} = K_{00} S'_{00} n_1 n_2 V_{12} \Delta t \quad (1)$$

where K_{00} is the neutral–neutral collision kernel, n_1 , n_2 and n_{12} are the particle number concentrations, and V_{12} is the van der Waals enhancement factor (see below). The neutral–neutral sticking probability for pure biogenic particles, $S'_{00,B}$, is:

$$S'_{00,B} = \exp[-0.693(C_B/m_{12})^{S_B}]$$

where C_B and S_B are free parameters. The parameter C_B effectively defines the threshold mass of stable clusters because the sticking probability $S'_{00,B} = 0.5$ when $C_B = m_{12}$, whereas the parameter S_B controls the sharpness of the threshold. The sticking probability for collisions where at least one particle is mainly sulfuric acid is similarly defined as:

$$S'_{00,A} = \exp[-0.693(C_A/m_{12})^{S_A}]$$

where C_A and S_A are free parameters.

The neutral–neutral collision kernel, K_{00} , in equation (1) is the Fuchs form of the Brownian coagulation coefficient^{47,48}. The van der Waals enhancement factor is the modification to Fuchs theory due to Scaats⁴⁹, as described in ref. 50, for a Knudsen number in the kinetic (free molecular) regime. The enhancement factor is:

$$V_{12} = 1 + \frac{\sqrt{A'/3}}{1 + b_0\sqrt{A'}} + b_1\ln(1 + A') + b_2\ln(1 + A')^3$$

where the reduced Hamaker constant, A' , is:

$$A' = \frac{A}{kT} \frac{r_1 r_2}{(r_1 + r_2)^2}$$

where $r_{1,2}$ are the particle radii, $A = 6.4 \times 10^{-20}$ J (the Hamaker constant for sulfuric acid⁵⁰), $b_0 = 0.0151$, $b_1 = -0.186$, $b_2 = -0.0163$, k is the Boltzmann constant and T is temperature. The same Hamaker constant is used for both sulfuric acid and HOMs because it does not noticeably change the model predictions.

Ions and charged clusters collide according to a similar expression as equation (1):

$$n_{12} = (E \times K_{00}) S' n_1 n_2 \Delta t \quad (2)$$

where E is an enhancement factor to obtain the charged collision kernels (described below). The sticking probability for collisions between a neutral particle and a charged particle, $S'_{0+,0-}$, is:

$$S'_{0+,0-} = \exp[-0.693(C/m_{12})^{S'_{0+,0-}}]$$

where $S'_{0+,0-}$ is a free parameter and $C = C_B$ or C_A for biogenic or acid particles, respectively. Ion–ion recombination results in a neutral particle, which may evaporate at small sizes. The model allows partial evaporation of such recombination particles; in this case the cluster divides into monomers and the mass is conserved. The probability of cluster survival after ion–ion recombination, S'_{+-} , is:

$$S'_{+-} = \exp[-0.693(C_{+-}/m_{12})^{S'_{+-}}]$$

where C_{+-} is a free parameter. A power of unity ($S'_{+-} = 1$) is used because the data do not constrain this parameter well.

To obtain the charged collision kernels, the neutral–neutral collision kernel is multiplied by size-dependent enhancement factors, E :

$$\begin{aligned} E'_{0+,0-} &= K_{0+,0-}/K_{00} \\ E_{+,+-} &= K_{+,+-}/K_{00} \\ E_{+-} &= K_{+-}/K_{00} \end{aligned}$$

where K are the collision kernels and the subscripts refer to the charge of the colliding particles. The charged collision kernels in equation (2) are obtained from ref. 51, which refers to sulfuric acid particles. Because biogenic particles may have different neutral–charged collision kernels, their enhancement factor is left free in the fit:

$$E_{0+,0-} = \frac{E'_{0+,0-} - 1}{f_{0+,0-}} + 1 \quad (3)$$

where $f_{0+,0-}$ is a free parameter.

Ions, monomers, clusters and larger particles are continually lost by diffusion to the walls and by dilution of the chamber contents with fresh gas mixture. The dilution lifetime is near 3 h (10^{-4} s⁻¹), depending on the total sampling rate of all instruments attached to the chamber. The wall loss rate is 1.8×10^{-3} s⁻¹ for H₂SO₄ monomers, and decreases with increasing cluster or molecule diameter as $1/d$. The same scaling law is used to obtain the wall loss rate for HOMs; that is, it is assumed that HOMs and particles that collide with the walls are irreversibly lost. For experimental runs for which there is a pre-existing population of particles in the chamber at the start of a run due to incomplete cleaning of the chamber, losses to this coagulation sink are accounted for by inserting the initial size distribution into the size bins of the model.

To determine the nucleation rates, the five free parameters of the model (S_B , S_A , $S_{0+,0-}$, $f_{0+,0-}$ and C_{+-}) are fitted to the experimental particle concentrations in the PSM1.8 and CPC2.5 versus time. For example, for neutral pure biogenic runs, only one free parameter (S_B) is involved in the fit. The value of S_B ranges from 12 to 14, S_A from 4 to 6, $S_{0+,0-}$ from 0.1 to 1.0, $f_{0+,0-}$ is near 4 and C_{+-} is near 10,000 Th. The parameters C_B , C_A , S_{+-} and f_c were determined by a global fit to all runs in the dataset and then subsequently fixed at these values. The fitted threshold masses for C_B and C_A are around 1,300 Th and 700 Th, respectively. The parameter S_{+-} is set to 1.0 and f_c is set to 0.52. The time development of the particle number

concentrations in both counters throughout all of the nucleation events in our dataset is well reproduced by the model (an example is shown in Extended Data Fig. 4b).

After fitting the data with the model, the nucleation rate $J_{1.7}$ is determined as the number of particles that grow to a mobility diameter of 1.7 nm or larger in any time-step, divided by the time increment. In each nucleation run at fixed conditions, the time $t_{\max}$ is determined at which $J_{1.7}$ is maximum; the value of $J_{1.7}$ for that run is then calculated as the mean measurement over the interval ($t_{\max} \pm 300$ s).

There are three major advantages of using a data-driven kinetic model to determine nucleation rates rather than making direct measurements with the PSM1.8 or CPC2.5 data. First, it avoids the need for time derivatives of the data, which are subject to large errors at low counting rates. Second, particle growth rates are determined by kinetics and properly account for growth due to collisions both with monomers and with other particles. The model treatment of the data therefore avoids the exponential sensitivity on experimental growth rates that occurs with other methods^{52–55}. Experimental growth rates are determined from particle counter rise times and have relatively large uncertainties in the 1–3-nm size range. Finally, the model requires consistency between the PSM1.8 and CPC2.5 so the formation rates are experimentally constrained both near the 1.7-nm threshold size and near 3 nm.

Verification of the model nucleation rates. We performed extensive cross-checks of the nucleation rates obtained with the model by calculating the nucleation rates independently in two additional ways: (1) direct measurements at 1.8 nm using the scanning PSM and (2) CPC2.5 measurements that are stepwise-corrected to 1.7-nm threshold size. Within their experimental uncertainties, the nucleation rates obtained by both these methods agree well with the values obtained with the AEROCLOUD kinetic model.

The stepwise-corrected method is described in detail in ref. 55, but a brief summary is provided here. The nucleation rates are derived from the rate of change of the formation rates, dN_{CPC}/dt , where N_{CPC} is the particle number concentration measured with the CPC2.5 above its detection threshold, d_{th} . The formation rate is corrected in two sequential steps for particle losses to chamber walls, dilution and coagulation: (1) particle losses above d_{th} and (2) particle losses during growth from 1.7 nm to d_{th} . The dilution and wall loss rates are the same as in the kinetic model. To calculate the coagulation rate, the particles are divided into size bins and then the loss rate in each bin i is computed by summing the size-dependent collision (coagulation-loss) rate of the particles in bin i with those in all other bins. The total coagulation loss rate is then the sum of the particle loss rates in each bin i .

Correcting for particle losses during growth from 1.7 nm to d_{th} (item (2) above) requires knowledge of the particle growth rate. This is experimentally determined with several instruments, for example, from the appearance times measured in the scanning PSM⁵⁶, which detects particles over a range of threshold diameters between 1 nm and 2.5 nm. The growth rates were also measured over different size ranges with several other instruments, including a fixed-threshold PSM, two DEGCPCs, a TSI 3776 CPC, an API-TOF, an NAIS, an nRDMA and a nano-SMPS. The experimental growth rates are parameterized because they cannot be measured sufficiently precisely at each point in time during all events. To determine the nucleation rate at 1.7 nm from the corrected formation rate at d_{th} , the size interval is divided into m log-normally spaced bins, $d\log(D_p)$, chosen to match the spacing of the SMPS bins at larger sizes. The residence time of a particle in each bin is $\delta t = \delta d_i / (\text{growth rate})$, where δd_i is the size of bin i . Starting with the measured particle distribution above d_{th} , the size distribution and formation rate is then extended towards 1.7 nm in a stepwise process. In the first step, using the known loss rates due to the chamber walls, dilution and coagulation, as well as the time δt , the concentration in the largest new bin is calculated, as well as the formation rate into this bin. Using this concentration, the size distribution is updated and the process is repeated until, after m steps, the smallest size bin at 1.7 nm is reached, where the nucleation rate is determined.

The NAIS. The neutral cluster and air ion spectrometer (NAIS)⁵⁷ measures the size distributions of positively and negatively charged particles, and also of total (charged plus neutral) particles, between mobility-equivalent diameters of 0.75 nm and 45 nm. Because the instrument includes two mobility analysers operating in parallel, positive and negative spectra are obtained simultaneously, each with 21 electrometers. Taking into account the internal diffusion losses, the mobility distribution is then calculated in 28 size bins from the measured electrometer currents.

The instrument operates sequentially in three modes: ion, particle and offset mode (one cycle takes 150 s). The aerosol sample first passes through a pre-conditioning section containing a discharger, an electric filter, a charger and a second electric filter (post-filter). The charger and discharger are corona needles of opposite polarities. In ion mode, the preconditioning unit is switched off and the sample passes through unaffected. In this way, the mobility analysers

measure only ions and charged particles from the CLOUD chamber. In particle mode—which was not used for the results reported here—both chargers are switched on and so neutral particles from the CLOUD chamber can be classified. The post-filters improve the measurements by removing residual ions from the charger. In offset mode, the dischargers and corresponding filters are switched on. The sample is charged to the opposite polarity as the subsequent analyser and so no detectable particles can enter. In this way, the noise levels and possible parasitic currents are measured to provide corrections for the preceding ion and particle measurement.

After preconditioning, the aerosol sample is classified in two cylindrical mobility analysers. The central electrode consists of several sections, each at a different fixed electric potential. The particles enter the analysers through a circular slit near the central electrode and are collected at the 21 outer electrodes where they transfer their charge to the connected electrometer and the resulting current is measured. The analysers operate at a sheath flow rate of 60 l min⁻¹. Filtered excess air serves as sheath gas to ensure conditions similar to the sample flow. The data inversion that converts the measured electrometer currents to particle concentrations is based on model calculations simulating trajectories of particles with different mobilities, and on calibration measurements of the internal losses. The performance of the NAIS for ion-mobility (size) and concentration measurements is described in refs 58, 59.

The API-TOF mass spectrometer. The atmospheric pressure interface time-of-flight (API-TOF) mass spectrometer¹⁴ measures the mass-to-charge ratio of positive or negative ions with an inlet at atmospheric pressure. The first stage of the instrument consists of an atmospheric pressure interface (API) section where ions are focused and guided by two quadrupoles and an ion lens through three chambers at progressively lower pressures down to 10⁻⁴ mbar. The second stage of the instrument is a time-of-flight (TOF) mass spectrometer at 10⁻⁶ mbar.

The API-TOF was connected to the CLOUD chamber via a 1" (21.7-mm inner diameter) sampling probe shared with the NAIS. A Y-splitter divided the total flow of 20 l min⁻¹ equally between the two instruments. The sample flow for the API-TOF was 0.8 l min⁻¹, with the remainder being discarded.

The API-TOF measurements were made during GCR and π^+ beam runs; that is, the ions were charged by GCRs or charged pions traversing the CLOUD chamber. Because the API-TOF can measure only one polarity at a time, positive and negative ions were measured in different runs. Different instrument settings were used during the campaigns to optimize detection in the low- or high-mass regions of the spectrum. The data were analysed with *tofTools*³⁵, developed by the University of Helsinki. The tool is implemented in MATLAB and allows complete processing of TOF data: averaging, mass calibration, baseline detection, peak fitting and high-resolution analysis.

The CI-API-TOF mass spectrometer. Two nitrate chemical ionization atmospheric pressure interface time-of-flight (CI-API-TOF) mass spectrometers were used to measure neutral sulfuric acid and HOMs. The instruments were operated by the University of Frankfurt (UFRA-CI) and the University of Helsinki (UHEL-CI); differences between the two instruments are indicated in this section by adding the UHEL-CI characteristics in parentheses after those of the UFRA-CI. The CI-API-TOF has been described previously^{36,37}. The sample air from the CLOUD chamber was drawn in through a 1/2" stainless steel tube at flow rate of 9 l min⁻¹ (10 l min⁻¹). An electrostatic filter was installed in front of each instrument to remove ions and charged clusters formed in the chamber. The geometry of both ion sources follows the design of ref. 60, but a corona charger³⁴ (X-ray generator) is used for ion generation. Dry air with nitric acid vapour is flushed over the ionizer to generate NO₃⁻ (HNO₃)_{j=0,2} ions. The ions are guided into the sample flow with an electric field, where they react with sulfuric acid and HOMs. The reaction time is approximately 50 ms (200 ms) before the ions enter the API section through a pinhole with a diameter of 350 μ m (300 μ m). The API section consists of three consecutive differentially pumped chambers where the pressure is progressively reduced and the ions are focused by two sets of quadrupoles and an ion lens system. The mass-to-charge ratios, m/z , of the ions that pass through these chambers are measured by a time-of-flight (TOF) mass spectrometer (Tofwerk AG).

The voltage settings in the API-TOF section influence the mass-dependent transmission efficiency. The transmission curves were determined in a series of calibration measurements in which various perfluorinated acid vapours of different m/z were passed into the instrument in sufficient amounts to saturate all the primary ions. In this way, a constant ion signal could be generated at each m/z and so the transmission efficiency could be determined relative to that of the primary ions mass range. The UFRA-CI operated at the same voltage settings for the entire data collection period; the UHEL-CI was operated in a switching mode between two voltage settings optimized for low and high m/z , respectively.

The raw data were analysed with the MATLAB *tofTools* package³⁵. The mass scale is calibrated to an accuracy of better than 10 p.p.m. using a two-parameter fit.

The concentration of sulfuric acid is calculated from the ratio of bisulfate ion counting rates (in s⁻¹) relative to primary ions as follows:

$$[\text{H}_2\text{SO}_4] = C \times \text{SL}_{\text{H}_2\text{SO}_4} \ln \left[1 + \frac{\text{HSO}_4^- + \text{HSO}_4^- \cdot \text{HNO}_3}{\sum_{j=0}^2 \text{NO}_3^- \cdot (\text{HNO}_3)_j} \right]$$

The factor $\text{SL}_{\text{H}_2\text{SO}_4}$ corrects for losses in the sampling line from the CLOUD chamber. The calibration coefficient, C , is determined by connecting the CI-API-TOF to a well-characterized H₂SO₄ generator⁶¹. The value of C depends on the voltage settings in the API-TOF section and was determined to be $6.5 \times 10^9 \text{ cm}^{-3}$ ($1.2 \times 10^{10} \text{ cm}^{-3}$ and $2.8 \times 10^9 \text{ cm}^{-3}$ for the high and low m/z settings, respectively), with an uncertainty of +50%/-33%. The H₂SO₄ detection limit is $5 \times 10^4 \text{ cm}^{-3}$ or slightly lower.

The concentration of a HOM at $m/z = i$ is calculated as follows:

$$[\text{HOM}] = CT_i \times \text{SL}_{E_1/E_2} \ln \left[1 + \frac{\text{HOM}_i \cdot \text{NO}_3^-}{\sum_{j=0}^2 \text{NO}_3^- \cdot (\text{HNO}_3)_j} \right]$$

Here, $\text{HOM}_i \cdot \text{NO}_3^-$ is the background-subtracted counting rate of the HOM. Background levels were measured by sampling air from the clean CLOUD chamber without any α -pinene present. The factor T_i is the mass-dependent transmission efficiency. The calibration coefficient, C , is the same as that obtained for sulfuric acid because HOMs and sulfuric acid were shown to have similar molecular collision rates with the nitrate ions¹⁶. Furthermore, the binding of NO₃⁻ with highly oxidized HOMs is found in the present study to be strong, so clustering should proceed at near the kinetic limit, as it does for NO₃⁻ with sulfuric acid. The factor SL_{E_1/E_2} corrects for losses in the sampling line from the CLOUD chamber. The values were determined for E_1 and E_2 separately, using experimentally determined diffusion coefficients, as $\text{SL}_{E_1} = 1.443$ and $\text{SL}_{E_2} = 1.372$.

The HOM monomers, E_1 , are the background-subtracted sum of the peaks in the m/z band 235–424 Th; the HOM dimers, E_2 , are the corresponding sum for 425–625 Th. Instrumental contamination peaks are excluded from the band summation, as are peaks assigned to the RO₂ radical (C₁₀H₁₅O_{6,8,10,12}, which correspond to $m/z = 293$ Th, 325 Th, 357 Th and 389 Th). Total HOMs is defined as the sum $\text{RO}_2 + E_1 + E_2$.

HOM yields. The HOM yields from either ozonolysis or OH· chemistry were calculated by assuming equal production and loss rates during steady-state¹⁶:

$$\frac{d[\text{HOM}]}{dt} = \gamma_{\text{Ox}} k_{\text{AP}+\text{Ox}} [\text{AP}] [\text{Ox}] - k_{\text{loss}} [\text{HOM}] = 0$$

where the yield, γ_{Ox} , is the fraction of α -pinene (AP) oxidation reactions leading to HOM formation, and 'Ox' signifies O₃ or OH·. The values of the rate constants (in cm³ per molecule per second) at 278 K for oxidation of α -pinene are $k_{\text{AP}+\text{O}_3} = 8.05 \times 10^{-17}$ and $k_{\text{AP}+\text{OH}} = 5.84 \times 10^{-11}$, from the International Union of Pure and Applied Chemistry (IUPAC)⁶² (the α -pinene + O₃ rate constant is updated on the IUPAC website at http://iupac.pole-ether.fr/htdocs/datasheets/pdf/Ox_VOC8_O3_apinene.pdf). The HOM wall loss rate was determined to be $1.1 \times 10^{-3} \text{ s}^{-1}$, assuming they are irreversibly lost. An additional loss is due to dilution of the chamber contents by makeup gases ($0.1 \times 10^{-3} \text{ s}^{-1}$). The total loss rates for HOMs is then $k_{\text{loss}} = 1.2 \times 10^{-3} \text{ s}^{-1}$.

During the experiments involving pure OH· chemistry, nitrous acid (HONO) concentrations ranging from 0.5 p.p.b.v. to 3 p.p.b.v. were photolysed by ultraviolet radiation from the fibre optic system to produce OH·. This led to a small contamination of NO in the chamber, which may potentially influence the HOM yield. The OH· concentrations in the CLOUD chamber were estimated using the PTR-TOF measurements of the difference of the α -pinene concentrations with no OH· present (ultraviolet off) and OH· present (ultraviolet on at different intensities). The decrease in α -pinene was due to only OH· reactions, because no O₃ was present in the chamber during these experiments. The accuracy for [OH·] is estimated to be $\pm 30\%$ (1 σ) including uncertainties in α -pinene measurements and reaction rate constant, which leads to a systematic scale uncertainty on the HOM production rate, $k_{\text{AP}+\text{OH}} [\text{AP}] [\text{OH}]$, of $\pm 40\%$ (1 σ). However, run-to-run uncertainties contribute substantially to the overall uncertainty as indicated by the error bars in Extended Data Fig. 2.

The SO₂-CIMS. The SO₂ chemical ionization mass spectrometer (SO₂-CIMS) uses CO₃⁻ primary ions to convert SO₂ to SO₅⁻, which is then measured in a quadrupole mass spectrometer with an API interface (Georgia Tech). The general design of the ion source is shown in ref. 60, but the primary ions are generated with a corona discharge³⁴. The corona needle holder was modified so that CO₂, O₂ and Ar are fed directly over the corona discharge. In this way, direct contact between the N₂ sheath flow and the discharge needle is avoided, which leads to a reduced

contamination by NO_3^- and maximizes the ratio of CO_3^- to NO_3^- . The reaction scheme for the ionization of SO_2 to SO_5^- can be found in ref. 63. The use of a dry N_2 buffer flow in front of the pinhole of the mass spectrometer evaporates associated water molecules from SO_5^- ions, and so sulfur dioxide is detected in the mass spectrum at $m/z = 112$ Th (SO_5^-).

The SO_2 concentration (in p.p.t.v.) is calculated from the ion count rates, $R_{m/z}$, as follows:

$$\text{SO}_2 = C_5 \ln(1 + R_{112}/R_{60})$$

where R_{112} corresponds to the background-corrected ion count rate of SO_5^- and R_{60} is the ion count rate of the primary ion CO_3^- . The calibration factor C_5 was obtained by periodically calibrating the instrument with a SO_2 gas standard (Carbagas AG) during the campaign. During a calibration, the gas standard was diluted with ultraclean humidified air at 38% relative humidity (the same as that supplied to the CLOUD chamber) to achieve a range of different SO_2 mixing ratios between 12 p.p.t.v. and 11 p.p.b.v. The calibration factor was found to be 1.3×10^5 p.p.t.v., with an estimated uncertainty of $\pm 11\%$. The error includes uncertainties in the flow rates during a calibration and in the gas standard concentration, as well as statistical uncertainties. However, we also observed that temperature changes in the experimental hall where the experiments were conducted led to a drift in the SO_5^- background signal when no SO_2 was applied to the CIMS. This effect contributes to the overall uncertainty and mainly affects the measurement at low SO_2 levels (< 100 p.p.t.v.), with lower precision in this concentration range. For example, at 30 p.p.t.v. SO_2 , the estimated uncertainty is $\pm 23\%$, but it becomes progressively smaller with higher SO_2 levels, reaching $\pm 13\%$ above 100 p.p.t.v. SO_2 . The detection limit of the instrument is 15 p.p.t.v. SO_2 .

Experimental errors. To determine $J_{1.7}$, the measured particle concentrations in the PSM1.8 and CPC2.5 versus time are fitted with the AEROCLOUD model (see above). The nucleation rate error, σ_J , has three main components. The dominant error at slow growth rates is due to uncertainties in the PSM1.8 and CPC2.5 detection thresholds for HOM particles⁶⁴. The threshold error components are first determined numerically for each nucleation measurement by performing additional AEROCLOUD fits after shifting the PSM1.8 particle detection threshold by $+0.2/-0.1$ nm and the CPC2.5 threshold by ± 0.4 nm. This provides four fractional $J_{1.7}$ errors which are then averaged for each counter to provide a mean fractional uncertainty, σ_{psm} and σ_{cpc} , respectively. The total error due to detection threshold uncertainties, σ_{thr} , for the combined fit to the PSM1.8 and CPC2.5 data is then:

$$\frac{1}{\sigma_{\text{thr}}^2} = \frac{1}{\sigma_{\text{psm}}^2} + \frac{1}{\sigma_{\text{cpc}}^2} + \frac{\sigma_{\text{psm}}\sigma_{\text{cpc}}}{(\sigma_{\text{psm}}^2 + \sigma_{\text{cpc}}^2)^{1/2}}$$

The total fractional $J_{1.7}$ error, σ_J , is then obtained by adding σ_{thr} in quadrature with an experimental error due to run-to-run reproducibility under nominally identical chamber conditions, σ_{exp} , and an error to account for model approximations, σ_{model} :

$$\sigma_J^2 = \sigma_{\text{thr}}^2 + \sigma_{\text{exp}}^2 + \sigma_{\text{model}}^2$$

where $\sigma_{\text{exp}} = 30\%$ and $\sigma_{\text{model}} = 50\%$.

The concentration of O_3 is measured with a calibrated instrument and is known to $\pm 10\%$. The α -pinene concentration in the PTR-TOF is known to $\pm 10\%$. As discussed above, the uncertainty on SO_2 is $\pm 13\%$ above 150 p.p.t.v., increasing at lower values to $\pm 23\%$ at 30 p.p.t.v.

For CI-API-TOF measurements, the run-to-run experimental uncertainties are $\pm 10\%$ for $[\text{H}_2\text{SO}_4]$ and $\pm 20\%$ for $[\text{HOM}]$. However, there is a larger overall systematic error that scales all measurements by the same amount. The systematic scale uncertainty for $[\text{H}_2\text{SO}_4]$ is estimated to be $+50\%/-33\%$. This estimate is based on a comparison of $[\text{H}_2\text{SO}_4]$ measurements with a CIMS and a calibrated H_2SO_4 generator⁶¹. The systematic uncertainties for $[\text{HOM}]$ have the following sources and fractional errors (1σ): sulfuric acid calibration (50%), charging efficiency of HOMs in the ion source (25%), mass dependent transmission efficiency (50%) and sampling line losses (20%). This results in an overall systematic scale uncertainty for $[\text{HOM}]$ of $+80\%/-45\%$. The uncertainty in the HOM yield from ozonolysis or hydroxyl chemistry is estimated by adding the $[\text{HOM}]$ uncertainty in quadrature with the errors for α -pinene (10%), O_3 (10%), OH· (30%), HOM wall loss rate (6%) and rate constants (35% for the α -pinene O_3 reaction and 20% for the α -pinene OH· reaction). This results in a mean estimated uncertainty in HOM yield for either ozonolysis or hydroxyl chemistry of $+100\%/-60\%$.

Quantum chemical calculations. To estimate the characteristic binding energies and evaporation rates expected for ELVOC clusters, we chose $\text{C}_{10}\text{H}_{14}\text{O}_7$ (molecular

weight of 246) to represent the ELVOC monomer, E_1 , and $\text{C}_{20}\text{H}_{30}\text{O}_{14}$ (molecular weight of 494) to represent the covalently bound ELVOC dimer, E_2 . Their formation mechanism and structures are shown in Extended Data Figs 3 and 7. To evaluate the effect of charge on the formation of ELVOC clusters, we studied initial molecular clusters of E_1 and E_2 that are either neutral or else include an ion of the type E_1^- , HSO_4^- , NO_3^- or NH_4^+ (Extended Data Table 1).

We calculated formation Gibbs free energies at 278 K, $\Delta G_{278\text{K}}$, of different clusters with the MO62X functional⁶⁵ and the 6-31+G(d) basis set⁶⁶ using the Gaussian09 program⁶⁷. The formation Gibbs free energy can be related to evaporation rate as described in refs 68, 69. In previous works^{10,15}, we used the method proposed in ref. 68 for calculating the formation free energy of different clusters. However, this method is too computationally demanding for the large clusters of the present study. The MO62X functional has been shown to be well suited to the study of atmospheric clusters⁷⁰. Ref. 70 has shown how reducing the basis set from the largest Pople basis set available (6-311++G(3df,3pd)) to the basis set used in this work (6-31+G(d)) leads to differences in the calculated formation free energies below 1 kcal mol^{-1} . Therefore, MO62X/6-31+(d) is a good alternative to the B3RCC2 method⁶⁸ when studying large clusters. We confirmed this by comparing the formation free energies previously calculated¹⁵ using the B3RCC2 method with those calculated here using the MO62X/6-31+G(d) method. The differences were found to be below 2 kcal mol^{-1} .

Parameterization of the pure biogenic nucleation rate. We parameterized the experimentally measured pure biogenic nucleation rates in a form suitable for global aerosol models. The neutral and ion-induced pure biogenic nucleation rates (in $\text{cm}^{-3} \text{ s}^{-1}$) are parameterized as:

$$J_n = a_1[\text{HOM}]^{a_2+a_5/[\text{HOM}]} \quad (4)$$

$$J_{\text{in}} = 2[n_{\pm}]a_3[\text{HOM}]^{a_4+a_5/[\text{HOM}]}$$

where $[n_{\pm}] = [n_+] = [n_-]$ is the small-ion concentration of either sign. Expressions for $[\text{HOM}]$ and $[n_{\pm}]$ are given in equations (7) and (10) below, respectively. The parameters a_n are determined from fits to the data in Fig. 3 and have the values $a_1 = 0.04001$, $a_2 = 1.848$, $a_3 = 0.001366$, $a_4 = 1.566$ and $a_5 = 0.1863$, with $[\text{HOM}]$ expressed in units of 10^7 cm^{-3} . The parameterized rates are shown by the curves in Fig. 3. The R^2 value of the fit is 0.97. The terms a_{1-4} describe simple power laws, whereas the term a_5 accounts for the steepening of the nucleation rate at low HOM concentrations. The nucleation rates are assumed to be independent of temperature, except for the effect of rate constants (equation (6) below), because the experimental measurements exist at only a single temperature.

The HOM concentration in equation (4) is determined from its production and loss rates:

$$\frac{d[\text{HOM}]}{dt} = Y_{\text{HOM}*\text{O}_3}k_{\text{MT}*\text{O}_3}[\text{MT}][\text{O}_3] + Y_{\text{HOM}*\text{OH}}k_{\text{MT}*\text{OH}}[\text{MT}][\text{OH}\cdot] - k_{\text{HOM}}[\text{HOM}] \quad (5)$$

where MT represents total monoterpenes. The IUPAC⁶² reaction rate constants (in cm^3 per molecule per second) for oxidation of α -pinene by ozone and hydroxyl radicals are, respectively:

$$k_{\text{MT}*\text{O}_3} = 8.05 \times 10^{-16} \exp(-640/T) \quad (6)$$

$$k_{\text{MT}*\text{OH}} = 1.2 \times 10^{-11} \exp(440/T)$$

where T (in K) is the temperature (the α -pinene+ O_3 rate constant is updated on the IUPAC website at http://iupac.pole-ether.fr/htdocs/datasheets/pdf/Ox_VOC8_O3_apinene.pdf). The HOM yields in each ozone-monoterpene and hydroxyl-monoterpene reaction are $Y_{\text{HOM}*\text{O}_3}$ and $Y_{\text{HOM}*\text{OH}}$, respectively. The parameter k_{HOM} is the HOM loss rate or, equivalently, the atmospheric condensation sink, CS (in s^{-1}). The condensation sink is determined assuming the diffusion characteristics of a typical α -pinene oxidation product (see appendix A1 of ref. 71). Assuming steady-state in equation (5), the HOM concentration becomes:

$$[\text{HOM}] = \frac{Y_{\text{HOM}*\text{O}_3}k_{\text{MT}*\text{O}_3}[\text{MT}][\text{O}_3] + Y_{\text{HOM}*\text{OH}}k_{\text{MT}*\text{OH}}[\text{MT}][\text{OH}\cdot]}{\text{CS}} \quad (7)$$

where the HOM yield from ozonolysis is $Y_{\text{HOM}*\text{O}_3} = 2.9\%$, and from reaction with the hydroxyl radical is $Y_{\text{HOM}*\text{OH}} = 1.2\%$ (Extended Data Fig. 2). The HOM yield from ozonolysis is determined from CLOUD measurements in the presence of a hydroxyl scavenger (0.1% H_2). The HOM yield from reaction with hydroxyl radicals is determined from CLOUD measurements in the absence of ozone, and where photolysed HONO provides the OH· source. Therefore, the experimental measurement of hydroxyl-initiated oxidation is made in the presence of NO_x , as occurs in the atmosphere.

The small-ion concentration in equation (4) is calculated from the steady-state solution of the ion balance equation:

$$\frac{d[n_{\pm}]}{dt} = q - \alpha[n_{\pm}]^2 - k_i[n_{\pm}] \quad (8)$$

where q (in $\text{cm}^{-3}\text{s}^{-1}$) is the ion-pair production rate and α is the ion-ion recombination coefficient (in cm^3s^{-1}). The factor of 2 in equation (4) accounts for nucleation from positive and negative ions. For the CLOUD GCR data, $q = 1.7\text{ cm}^{-3}\text{s}^{-1}$. Terrestrial radioactivity such as radon contributes additional ionization in the boundary layer over land masses⁷². The ion loss rate, k_i , is due to the condensation sink, CS, and ion-induced nucleation:

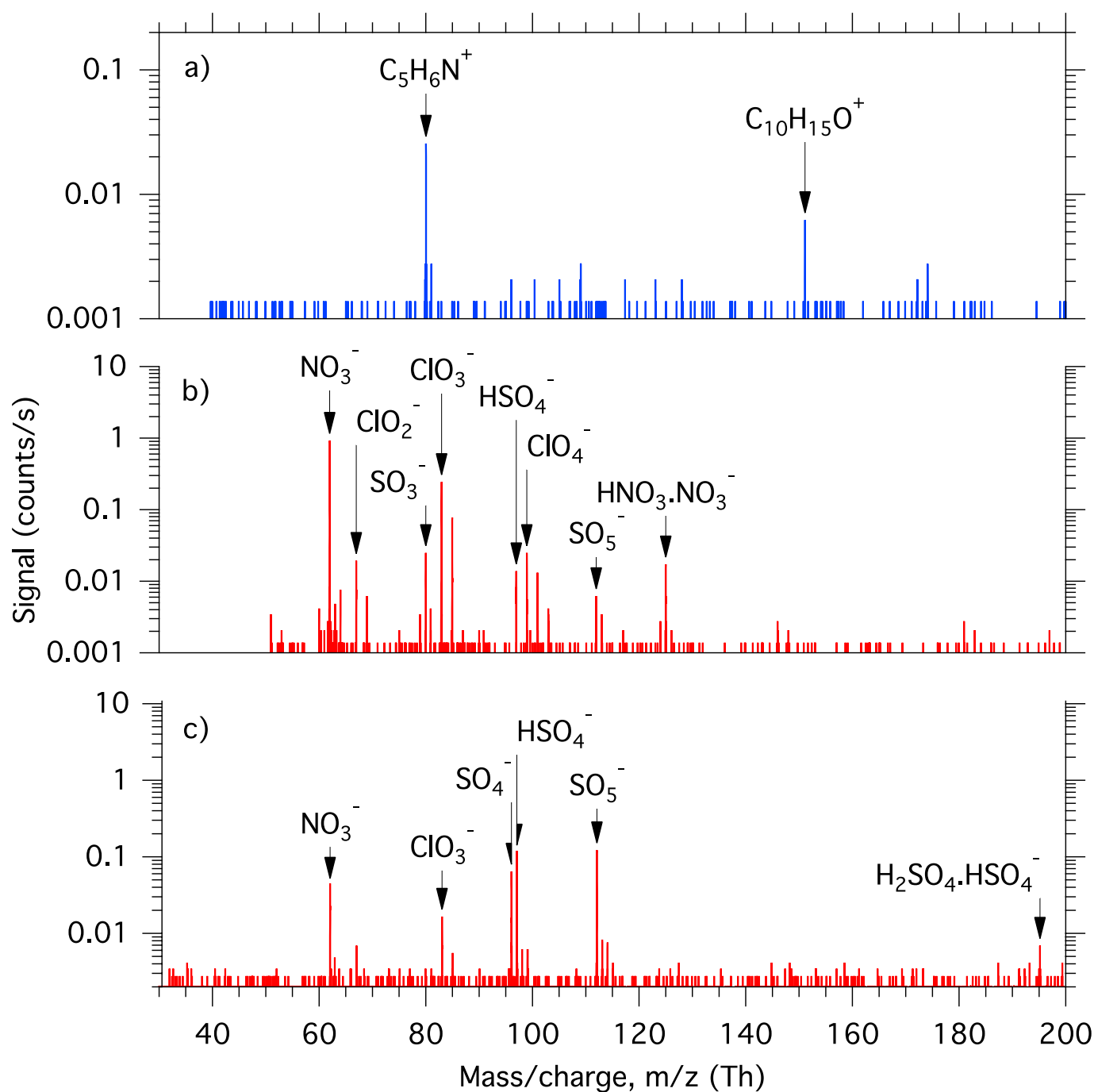
$$k_i = \text{CS} + \frac{J_{\text{in}}}{2[n_{\pm}]} \quad (9)$$

where $J_{\text{in}}/(2[n_{\pm}])$ is given by equation (4) and the steady-state concentration of small ions is, from equation (8):

$$[n_{\pm}] = \frac{(k_i^2 + 4\alpha q)^{1/2} - k_i}{2\alpha} \quad (10)$$

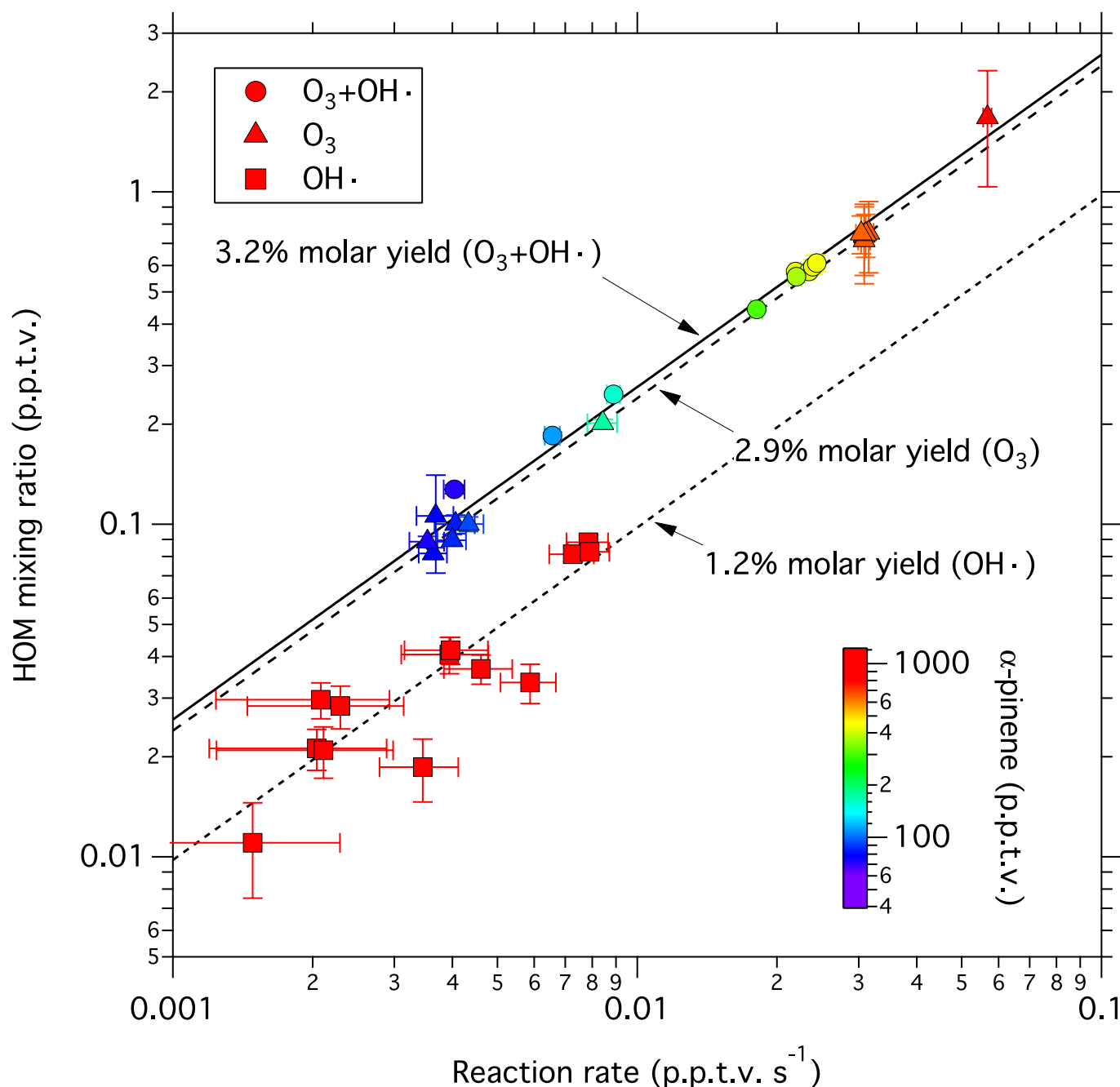
From equations (8) and (9), J_{in} saturates at $2q$ at high nucleation rates (see Fig. 3).

33. Enghoff, M. B., Pedersen, J. O. P., Uggerhøj, U. I., Paling, S. M. & Svensmark, H. Aerosol nucleation induced by a high energy particle beam. *Geophys. Res. Lett.* **38**, L09805 (2011).
34. Kürten, A., Rondo, L., Ehrhart, S. & Curtius, J. Performance of a corona ion source for measurement of sulphuric acid by chemical ionisation mass spectrometry. *Atmos. Meas. Tech.* **4**, 437–443 (2011).
35. Junninen, H. *et al.* A high-resolution mass spectrometer to measure atmospheric ion composition. *Atmos. Meas. Tech.* **3**, 1039–1053 (2010).
36. Jokinen, T. *et al.* Atmospheric sulphuric acid and neutral cluster measurements using CI-API-TOF. *Atmos. Chem. Phys.* **12**, 4117–4125 (2012).
37. Kürten, A. *et al.* Neutral molecular cluster formation of sulfuric acid-dimethylamine observed in real time under atmospheric conditions. *Proc. Natl Acad. Sci. USA* **111**, 15019–15024 (2014).
38. Graus, M., Müller, M. & Hansel, A. High resolution PTR-TOF: quantification and formula confirmation of VOC in real time. *J. Am. Soc. Mass Spectrom.* **21**, 1037–1044 (2010).
39. Mirme, S. *et al.* Atmospheric sub-3 nm particles at high altitude. *Atmos. Chem. Phys.* **10**, 437–451 (2010).
40. Zhang, S. H., Akutsu, Y., Russell, L. M., Flagan, R. C. & Seinfeld, J. H. Radial differential mobility analyzer. *Aerosol Sci. Technol.* **23**, 357–372 (1995).
41. Vanhanen, J. *et al.* Particle size magnifier for nano-CN detection. *Aerosol Sci. Technol.* **45**, 533–542 (2011).
42. Iida, K., Stolzenburg, M. R. & McMurry, P. H. Effect of working fluid on sub-2 nm particle detection with a laminar flow ultrafine condensation particle counter. *Aerosol Sci. Technol.* **43**, 81–96 (2009).
43. Wimmer, D. *et al.* Performance of diethylene glycol-based particle counters in the sub-3 nm size range. *Atmos. Meas. Tech.* **6**, 1793–1804 (2013).
44. Taira, M. & Kanda, Y. Continuous generation system for low-concentration gaseous nitrous acid. *Anal. Chem.* **62**, 630–633 (1990).
45. Heland, J., Kleffmann, J., Kurtenbach, R. & Wiesen, P. A new instrument to measure gaseous nitrous acid (HONO) in the atmosphere. *Environ. Sci. Technol.* **35**, 3207–3212 (2001).
46. Riccobono, F. *et al.* Contribution of sulfuric acid and oxidized organic compounds to particle formation and growth. *Atmos. Chem. Phys.* **12**, 9427–9439 (2012).
47. Fuchs, N. A. *The Mechanics of Aerosols* (Pergamon, 1964).
48. Seinfeld, J. H. & Pandis, S. N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* 2nd edn, 600 (Wiley, 2006).
49. Scaets, M. G. Brownian coagulation in aerosols—the role of long range forces. *J. Coll. Interf. Sci.* **129**, 105–112 (1989).
50. Chan, T. W. & Mozurkewich, M. Measurement of the coagulation rate constant for sulfuric acid particles as a function of particle size using tandem differential mobility analysis. *J. Aerosol Sci.* **32**, 321–339 (2001).
51. Laakso, L. *et al.* Kinetic nucleation and ions in boreal forest particle formation events. *Atmos. Chem. Phys.* **4**, 2353–2366 (2004).
52. Kerminen, V.-M. & Kulmala, M. Analytical formulae connecting the “real” and the “apparent” nucleation rate and the nuclei number concentration for atmospheric nucleation events. *J. Aerosol Sci.* **33**, 609–622 (2002).
53. Kulmala, M. & Kerminen, V.-M. On the formation and growth of atmospheric nanoparticles. *Atmos. Res.* **90**, 132–150 (2008).
54. Ehrhart, S. & Curtius, J. Influence of aerosol lifetime on the interpretation of nucleation experiments with respect to the first nucleation theorem. *Atmos. Chem. Phys.* **13**, 11465–11471 (2013).
55. Kürten, A., Williamson, C., Almeida, J., Kirkby, J. & Curtius, J. On the derivation of particle nucleation rates from experimental formation rates. *Atmos. Chem. Phys.* **15**, 4063–4075 (2015).
56. Lehtipalo, K. *et al.* Methods for determining particle size distribution and growth rates between 1 and 3 nm using the Particle Size Magnifier. *Bor. Environ. Res.* **19** (Suppl. B), 215–236 (2014).
57. Mirme, S. & Mirme, A. The mathematical principles and design of the NAIS – a spectrometer for the measurement of cluster ion and nanometer aerosol size distributions. *Atmos. Meas. Tech.* **6**, 1061–1071 (2013).
58. Asmi, E. *et al.* Results of the first air ion spectrometer calibration and intercomparison workshop. *Atmos. Chem. Phys.* **9**, 141–154 (2009).
59. Gagné, S. *et al.* Intercomparison of air ion spectrometers: an evaluation of results in varying conditions. *Atmos. Meas. Tech.* **4**, 805–822 (2011).
60. Eisele, F. L. & Tanner, D. J. Measurement of the gas-phase concentration of H_2SO_4 and methane sulfonic acid and estimates of H_2SO_4 production and loss in the atmosphere. *J. Geophys. Res. Atmos.* **98**, 9001–9010 (1993).
61. Kürten, A., Rondo, L., Ehrhart, S. & Curtius, J. Calibration of a chemical ionization mass spectrometer for the measurement of gaseous sulphuric acid. *J. Phys. Chem. A* **116**, 6375–6386 (2012).
62. Atkinson, R. *et al.* Evaluated kinetic and photochemical data for atmospheric chemistry: volume II – gas phase reactions of organic species. *Atmos. Chem. Phys.* **6**, 3625–4055 (2006).
63. Möhler, O., Reiner, T. H. & Arnold, F. The formation of SO_4^- by gas-phase ion-molecule reactions. *J. Chem. Phys.* **97**, 8233–8239 (1992).
64. Kangasluoma, J. *et al.* Sub-3 nm particle size and composition dependent response of a nano-CPC battery. *Atmos. Meas. Tech.* **7**, 689–700 (2014).
65. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120**, 215–241 (2008).
66. Ditchfield, R., Hehre, W. J. & Pople, J. A. Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **54**, 724–728 (1971).
67. Frisch, M. J. *et al.* Gaussian 09 Revision A.01, http://www.gaussian.com/g_prod/g09.htm (Gaussian, Inc., 2009).
68. Ortega, I. K. *et al.* From quantum chemical formation free energies to evaporation rates. *Atmos. Chem. Phys.* **12**, 225–235 (2012).
69. Ortega, I. K. *et al.* Can highly oxidized organics contribute to atmospheric new particle formation? *J. Phys. Chem. A* **120**, 1452–1458 (2016).
70. Elm, J. & Mikkelsen, K. V. Computational approaches for efficient modelling of small atmospheric clusters. *Chem. Phys. Lett.* **615**, 26–29 (2014).
71. Mann, G. W. *et al.* Intercomparison of modal and sectional aerosol microphysics representations within the same 3-D global chemical transport model. *Atmos. Chem. Phys.* **12**, 4449–4476 (2012).
72. Zhang, K. *et al.* Radon activity in the lower troposphere and its impact on ionization rate: a global estimate using different radon emissions. *Atmos. Chem. Phys.* **11**, 7817–7838 (2011).
73. Rissanen, M. P. *et al.* The formation of highly oxidized multifunctional products in the ozonolysis of cyclohexene. *J. Am. Chem. Soc.* **136**, 15596–15606 (2014).
74. Rissanen, M. P. *et al.* Effects of chemical complexity on the autoxidation mechanisms of endocyclic alkene ozonolysis products: from methylcyclohexenes toward understanding α -pinene. *J. Phys. Chem. A* **119**, 4633–4650 (2015).



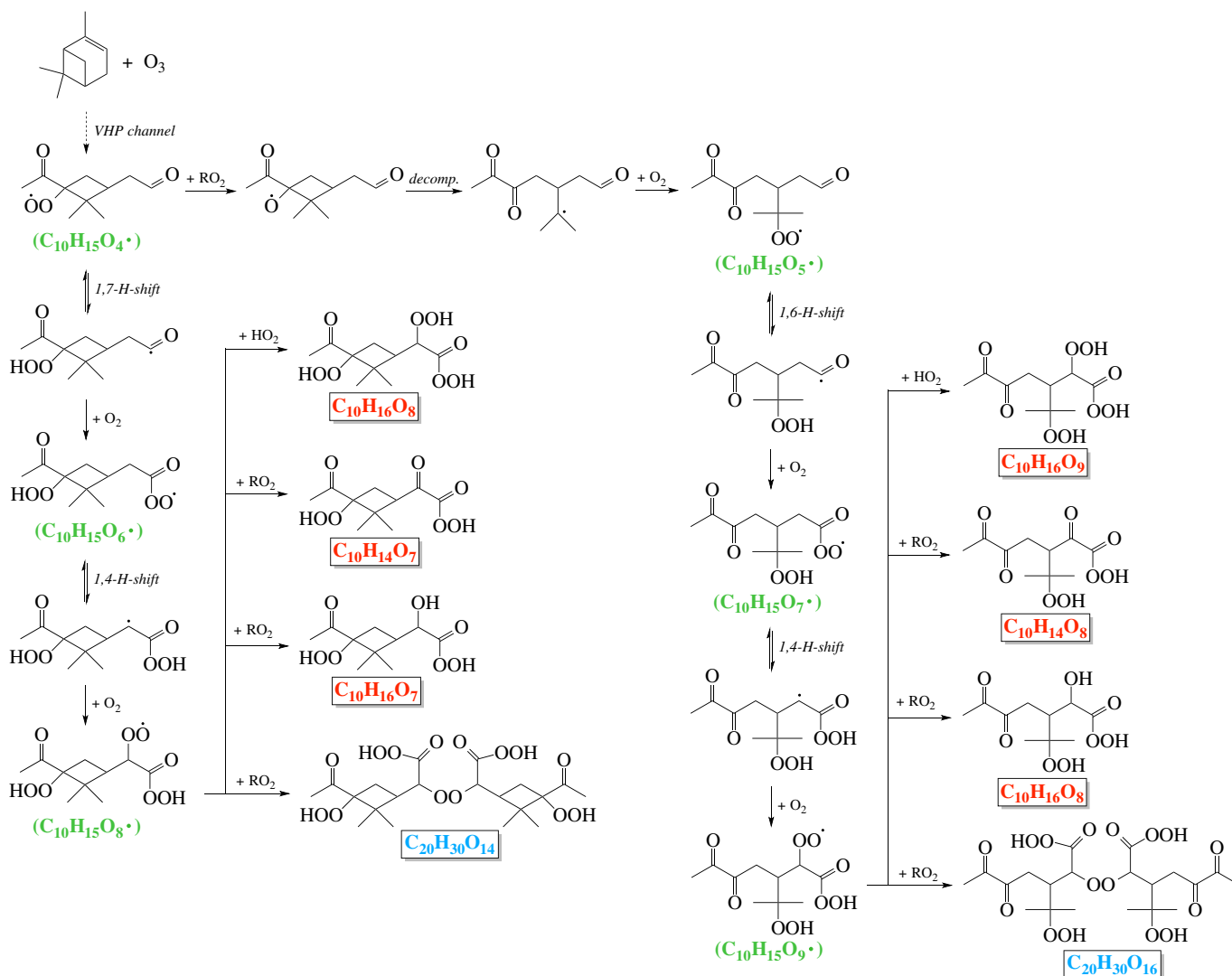
Extended Data Figure 1 | Small-ion mass spectra. **a, b,** Composition of positive (**a**) and negative (**b**) small ions measured by the API-TOF under GCR conditions and before adding any SO_2 to the chamber. The experimental conditions are zero α -pinene, 35 p.p.b.v. O_3 , zero H_2 or HONO, 38% relative humidity, 278 K and $[\text{H}_2\text{SO}_4] < 5 \times 10^4 \text{ cm}^{-3}$. Collisions will transfer positive charge to contaminant molecules having the highest proton affinity (**a**), and negative charge to contaminant molecules with the lowest proton affinity, that is, highest gas-phase acidity (**b**). From molecular cluster measurements, the positive ions

also include ammonium (NH_4^+), but its mass is below the set acceptance cut-off of the API-TOF. **c,** The negative small-ion spectrum at $[\text{H}_2\text{SO}_4] = 1.2 \times 10^5 \text{ cm}^{-3}$, after adding 32 p.p.t.v. SO_2 to the chamber, showing that the dominant ions species shift from nitrate to sulfur-containing. The experimental conditions are 340 p.p.t.v. α -pinene, 35 p.p.b.v. O_3 , zero H_2 or HONO, 38% relative humidity and 278 K. Water molecules evaporate rapidly from most hydrated ions in the API-TOF and so are not detected.



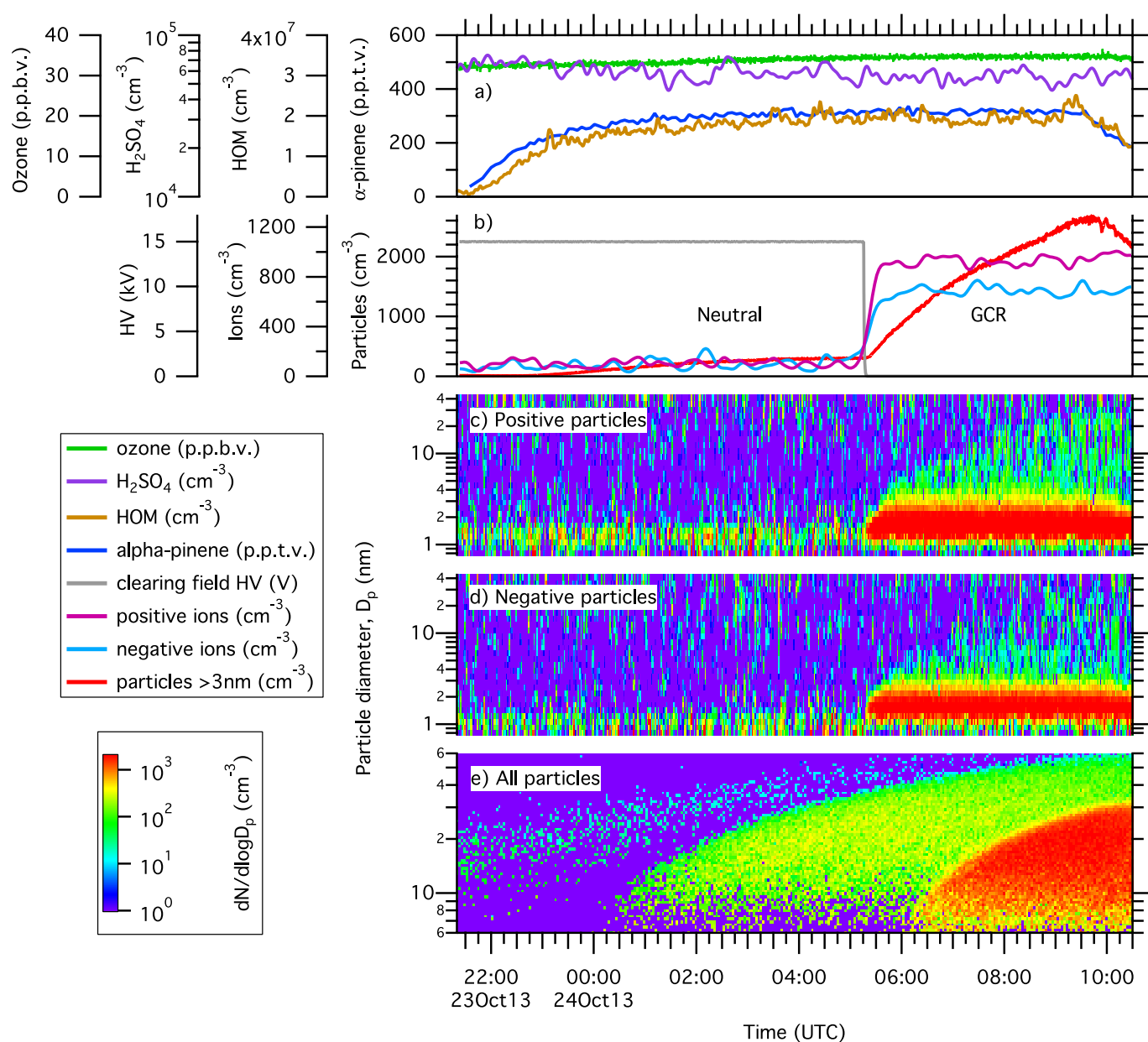
Extended Data Figure 2 | HOM yields versus α -pinene oxidation rates with O₃ and OH. Total HOM mixing ratios versus α -pinene reaction rate with (i) O₃ plus OH· (ozone without H₂ scavenger; circles and solid line), (ii) O₃ alone (ozone with 0.1% H₂ scavenger; triangles and dashed line) and (iii) OH· alone (produced by ultraviolet photolysis of nitrous acid, HONO, in the absence of O₃; squares and dotted line). The yields are shown for total HOMs = RO₂· + E₁ + E₂. The experimental conditions are 38% relative humidity, 278 K and (i) 70–440 p.p.t.v. α -pinene,

21–35 p.p.b.v. O₃, zero H₂ or HONO, 0%–100% ultraviolet, (ii) 80–1,230 p.p.t.v. α -pinene, 21–35 p.p.b.v. O₃, 0.1% H₂, zero HONO, 0%–100% ultraviolet, and (iii) 840–910 p.p.t.v. α -pinene, zero O₃ or H₂, 0.5–3 p.p.b.v. HONO, 0%–100% ultraviolet. The bars indicate 1σ point-to-point errors. Overall systematic scale uncertainties of ±40% for the reaction rates and +80%/–45% for the HOM mixing ratios are not shown. The combined errors on the HOM molar yields for either ozonolysis or hydroxyl chemistry are +100%/–60% (±1σ).



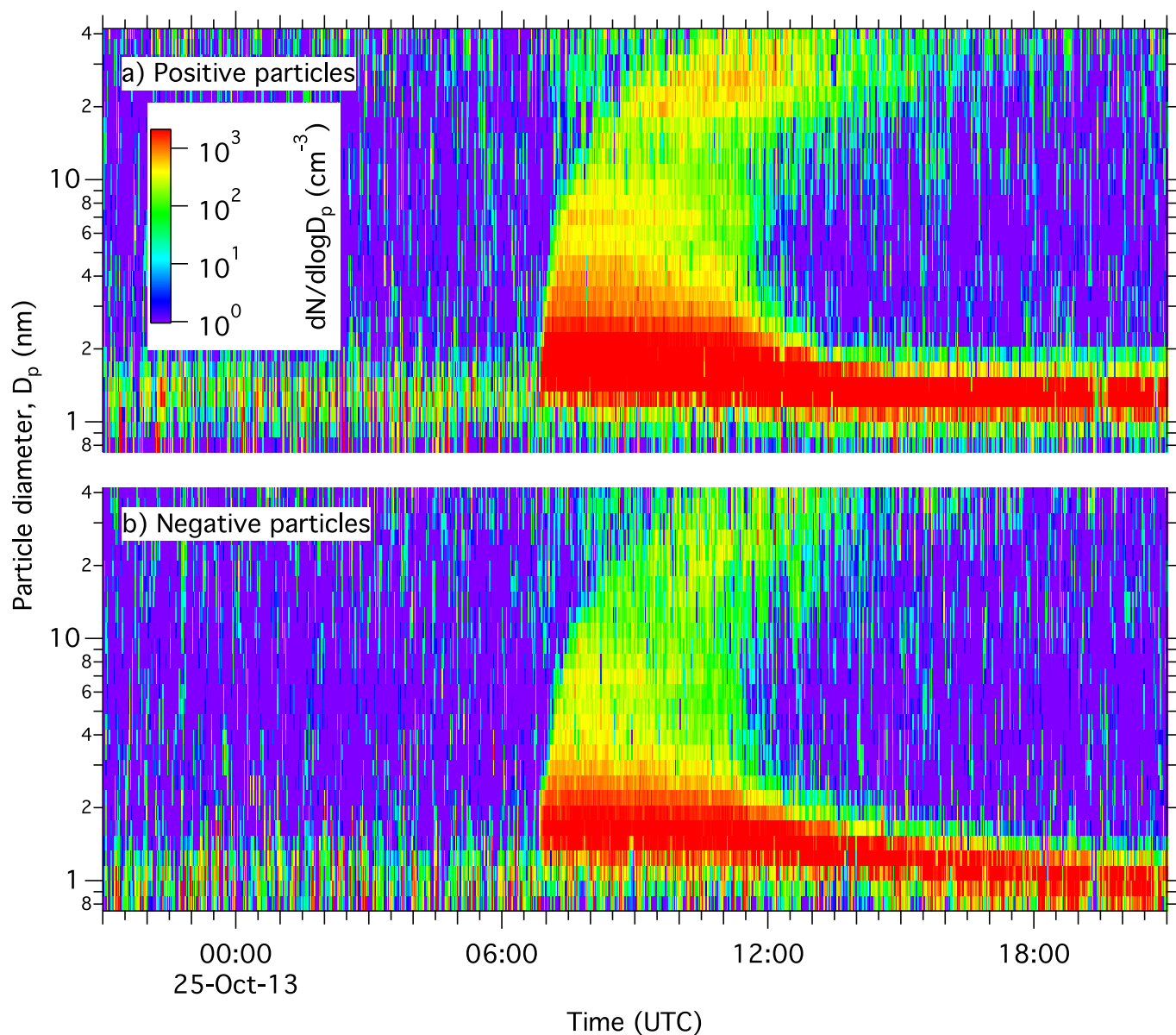
Extended Data Figure 3 | Proposed mechanism for the formation of the E₁ and E₂ surrogates via peroxy radical formation. The proposed scheme for the formation of the ELVOC monomer ($C_{10}H_{14}O_7$) and dimer ($C_{20}H_{30}O_{14}$) surrogates selected for quantum chemical calculations (Extended Data Fig. 7) is based on recently established autooxidation mechanisms for a series of cycloalkane + O_3 systems^{16,18–20,73,74}. Peroxy radicals in the figure are indicated by a green label, E₁ by a red label and E₂ by a blue label. Addition of ozone to the double bond of α -pinene produces two carbonyl-substituted Criegee biradicals. The energy-rich Criegee biradical is either collisionally stabilized, or isomerizes via 1,4-H-shift to a vinylhydroperoxide (VHP), which then decomposes to yield an OH· and an alkenoxy radical. The alkenoxy radical reacts with O_2 , leading to a peroxy radical, which is the potential precursor to a sequence of autooxidation reactions leading to the formation of HOMs¹⁶. Here the peroxy radical $C_{10}H_{15}O_4^\bullet$ is chosen as the starting point for HOM formation. The first intramolecular hydrogen abstraction is likely to take place at the aldehydic carbon from the opposite side of the peroxy group, although the rigid four-carbon-atom ring could hinder bending of the structure. For the *cis* configuration where the peroxy group and the aldehydic hydrogen are on the same side of the cyclobutyl ring, the 1,7-H shift rate is calculated²⁰ to be 0.14 s^{-1} , which initiates the autooxidation chemistry on a fast timescale compared to the HOM lifetime resulting from loss to the CLOUD chamber walls (about 900 s). The resultant acyclic radical undergoes rapid O_2 addition, leading to an -OOH functionalized peroxyacyl radical ($C_{10}H_{15}O_6^\bullet$). The second intramolecular hydrogen abstraction is expected to proceed at the carbon atom in the α position of the peroxyacyl group via 1,4-H isomerization. The resultant $C_{10}H_{15}O_8^\bullet$

terminates by known reactions of peroxy radicals ($HO_2^\bullet$ or $RO_2^\bullet$ under the present experimental conditions), producing a spectrum of HOM monomers that includes the E₁ surrogate, $C_{10}H_{14}O_7$. The homogeneous recombination of two peroxy radicals via elimination of O_2 produces the covalently bound dimer $C_{20}H_{30}O_{14}$ chosen as the E₂ surrogate. Alternatively, $C_{10}H_{15}O_8^\bullet$ can undergo further autooxidation, if sufficiently labile hydrogen atoms are available, leading to the observed closed-shell monomers with $\geq 9\text{ O}$ (Fig. 1). The self/cross-reaction of the $C_{10}H_{15}O_4^\bullet$ peroxy radical produces an alkoxy radical, which decomposes rapidly, leading to a carbonyl-functionalized peroxy radical ($C_{10}H_{15}O_5^\bullet$). This peroxy radical is another potential starting structure for HOM formation. The carbon-ring-opening reaction pathway, while increasing the steric availability of the H atom, might be a slow step. The effective formation rate of the C=O-functionalized peroxy radical is calculated to be less than about 10^{-3} s^{-1} , which is comparable to its wall deposition rate. The timescale with respect to the subsequent autooxidation reaction, on the other hand, is expected to be of the order of seconds, by analogy with that for branched-chain peroxy radicals⁷³. The unbalanced sources and sinks potentially account for the low signals of peroxy radicals with odd oxygen numbers (for example, $C_{10}H_{15}O_5^\bullet$, $C_{10}H_{15}O_7^\bullet$ and $C_{10}H_{15}O_9^\bullet$). The autooxidation process of $C_{10}H_{15}O_5^\bullet$ is presumed to proceed by an autooxidative reaction pathway similar to that for $C_{10}H_{15}O_4^\bullet$, eventually leading to the spectrum of HOM monomers and dimers observed in the CLOUD chamber. Except for the autooxidation channel, all the peroxy radicals are still subject to well-established reactions such as $R(\cdot)O_2 + RO_2/HO_2$, which are potentially important if the reaction rate is comparable to that for the autooxidation channel.



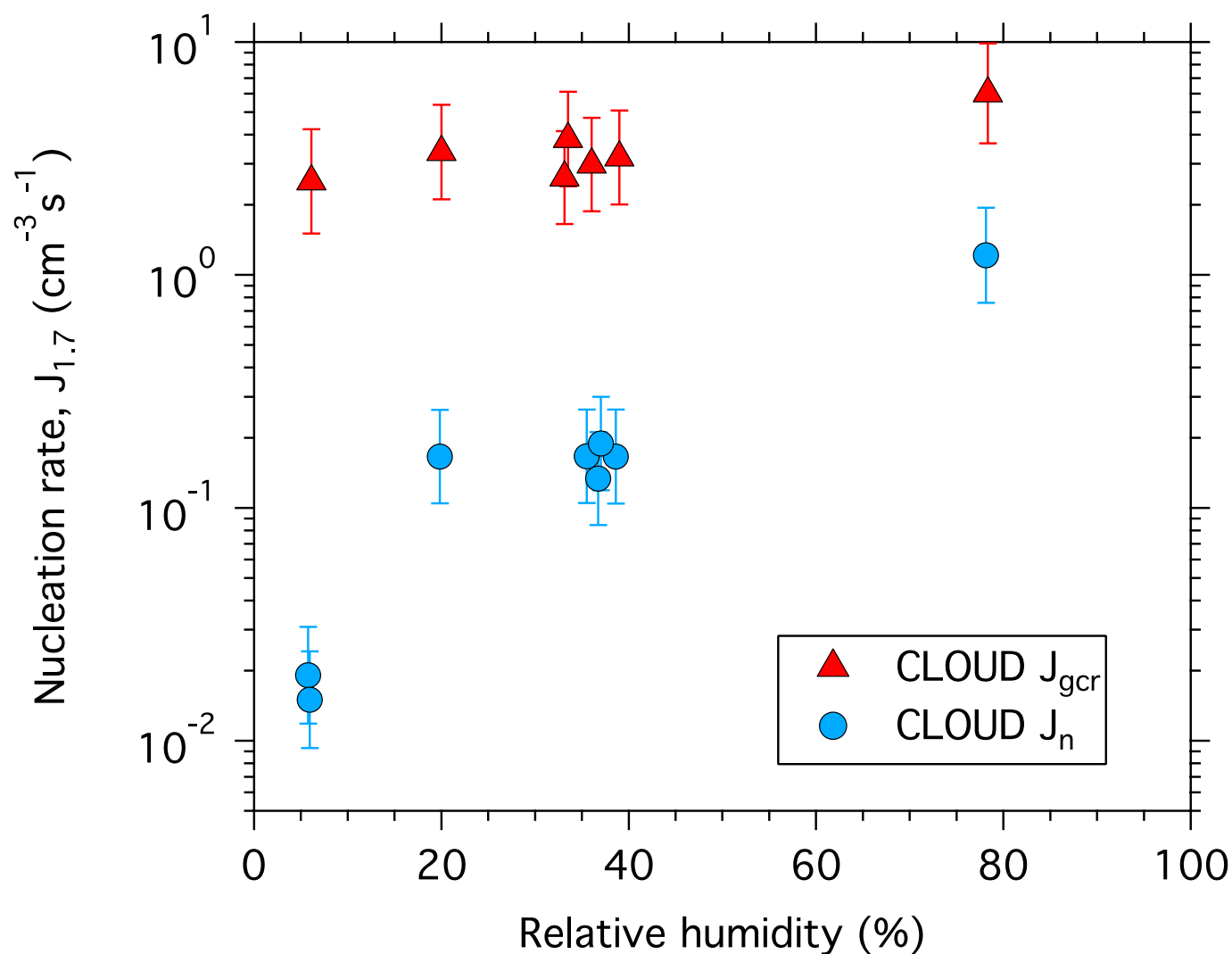
Extended Data Figure 4 | Typical nucleation run sequence. Example of a typical measurement sequence of the neutral and GCR nucleation rates as a function of coordinated universal time (UTC), at zero H_2 or HONO, 38% relative humidity and 278 K. **a**, The run began at 21:22, 23 October 2013, by starting the α -pinene flow into the chamber to reach a chosen equilibrium value near 300 p.p.t.v., which produced an equilibrium total HOMs concentration near $2 \times 10^7 \text{ cm}^{-3}$ (0.8 p.p.t.v.). **b**, Particles (red curve) formed at a slow rate in the chamber without ions present ('neutral' conditions). The clearing field high voltage (HV) was turned off at 05:16, 24 October 2013, and the subsequent presence of ions in the chamber from GCRs caused a sharp increase in the particle formation rate by about one order of magnitude (as seen by the increase in the gradient of the red curve). The nucleation rates are measured under constant gas conditions in the period before ($J_n = 0.14 \text{ cm}^{-3} \text{ s}^{-1}$) and after ($J_{\text{GCR}} = 3.3 \text{ cm}^{-3} \text{ s}^{-1}$) turning off the clearing field high voltage. **c**, **d**, Ion-induced nucleation is

observed both for positive (**c**) and negative (**d**) charged particles, followed by rapid particle growth to sizes above 10 nm. **e**, The nucleated particles grew over a period of several hours to diameters approaching 50 nm, where they begin to constitute cloud condensation nuclei. A sharp increase in the formation rate of particles above the SMPS detection threshold of 5 nm can be seen when GCR ions are present. The colour scale in **c–e** indicates $dN/d\log(D_p)$, where N (in cm^{-3}) is the particle number concentration and D_p (in nm) is the particle diameter. The concentrations of ozone and contaminant H_2SO_4 were essentially constant during the run, which ended at 09:30 when the α -pinene flow to the chamber was turned off. The H_2SO_4 measurement near $5 \times 10^4 \text{ cm}^{-3}$ corresponds to the instrumental background level of the CI-API-TOF mass spectrometer and so represents an upper limit on the actual concentration. Further characteristics of this run can be seen in Fig. 1.



Extended Data Figure 5 | Ion-induced nucleation event without H_2SO_4 , measured in the NAIS. a, b, Example of a nucleation event showing the growth versus time of positive (a) and negative (b) charged particles at 530 p.p.t.v. α -pinene, 35 p.p.b.v. O_3 , zero H_2 or HONO, $3.4 \times 10^7 \text{ cm}^{-3}$ HOM, 38% relative humidity, 278 K and $[\text{H}_2\text{SO}_4] < 5 \times 10^4 \text{ cm}^{-3}$. The colour scale shows the concentration of ions and charged particles. The clearing field high voltage was turned off at 06:48, marking the start of

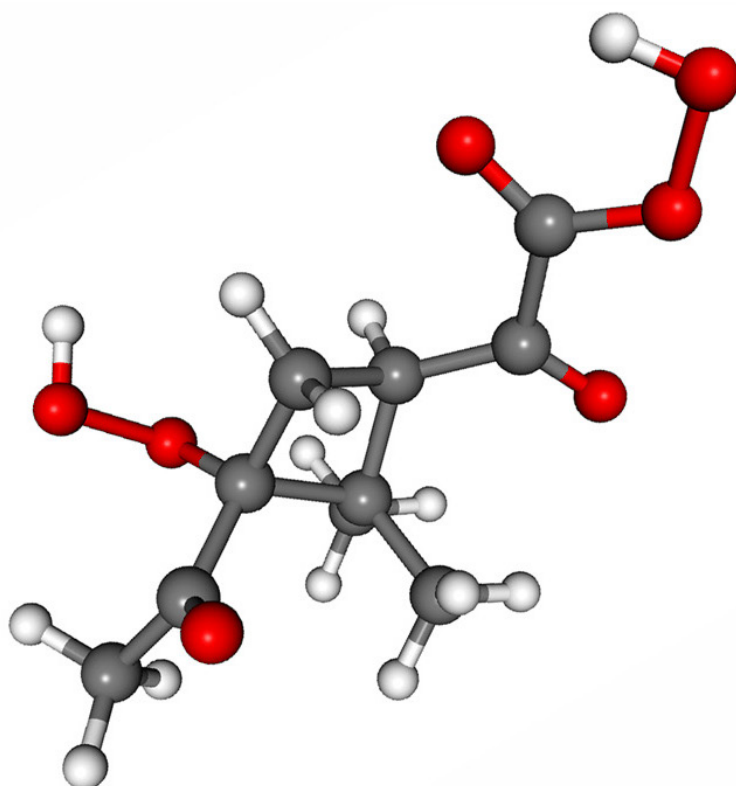
GCR ionization conditions in the chamber, and the α -pinene flow into the chamber was stopped at 10:52. Ion-induced nucleation can be seen for positive and negative charged particles, followed by rapid growth to sizes above 10 nm. Ion-ion recombination progressively neutralizes the charged particles as they grow, but some reappear at larger sizes, owing to diffusion charging.



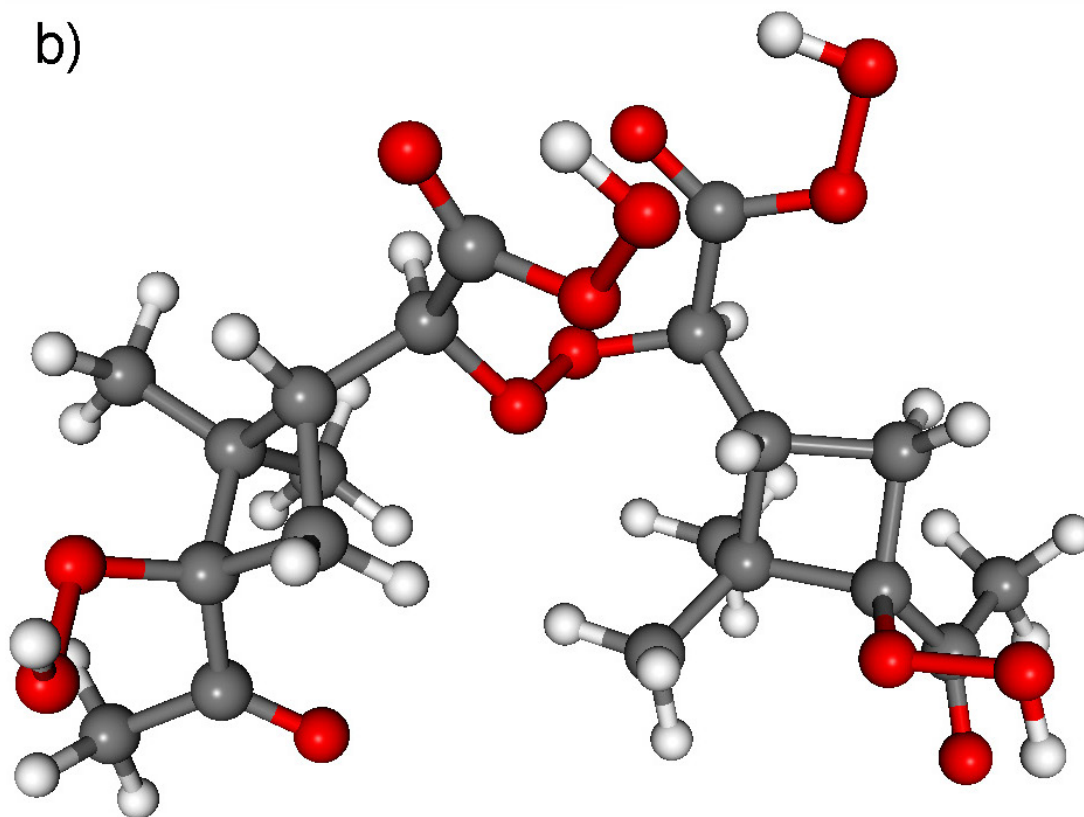
Extended Data Figure 6 | Nucleation rates versus relative humidity. Neutral (J_n ; circles) and GCR (J_{gcr} ; triangles) nucleation rates versus relative humidity. The experimental conditions are 250–800 p.p.t.v. α -pinene, 30–35 p.p.b.v. O_3 , zero H_2 or HONO, $(1.1–2.9) \times 10^7 \text{ cm}^{-3}$ HOM, 278 K and $(0.5–1.5) \times 10^5 \text{ cm}^{-3}$ H_2SO_4 . All measurements have

been corrected to the same total HOMs concentration ($2.05 \times 10^7 \text{ cm}^{-3}$) using the curves shown in Fig. 3. The bars indicate 1σ total errors, although these are not shown in the x direction because they are smaller than the symbols.

a)

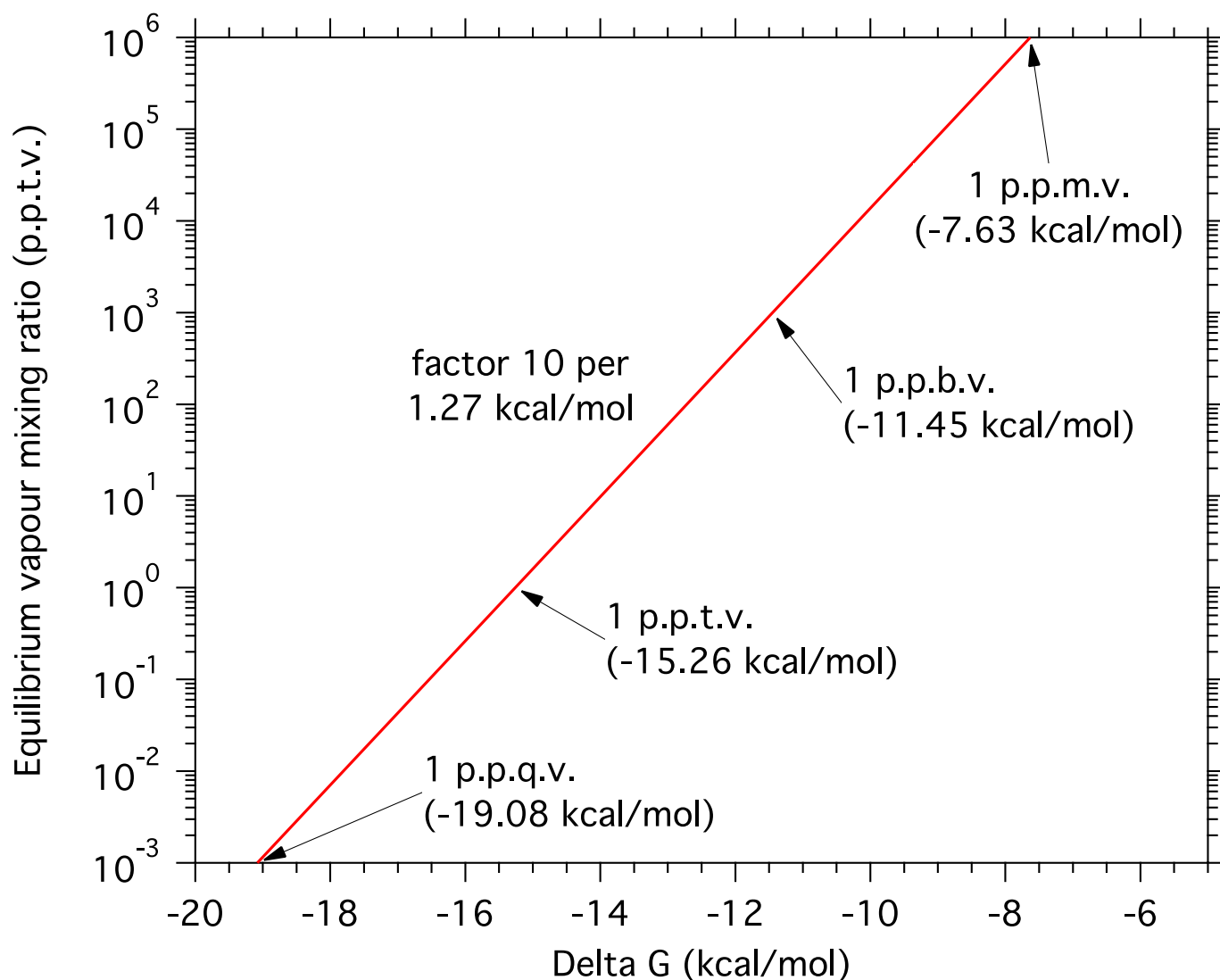


b)



Extended Data Figure 7 | Surrogate molecules chosen for quantum chemical calculations. **a, b,** Structures of the surrogate molecules chosen for quantum chemical calculations to represent the ELVOC monomer, E_1 , $C_{10}H_{14}O_7$ (**a**) and the covalently bound dimer, E_2 , $C_{20}H_{30}O_{14}$ (**b**).

Grey spheres represent carbon atoms, red are oxygen atoms and white are hydrogen atoms. We show their proposed formation mechanisms in Extended Data Fig. 3.



Extended Data Figure 8 | Relationship between cluster formation energies and equilibrium evaporation/condensation rates.

Estimated ELVOC vapour mixing ratios versus the $\Delta G_{278\text{ K}}$ at which the condensation and evaporation rates of the cluster at 278 K are in equilibrium^{68,69}. For example, a formation free energy of $-15.3\text{ kcal mol}^{-1}$

corresponds to equal rates for particle evaporation and vapour condensation at 278 K and 1 p.p.t.v. ELVOC vapour mixing ratio ($2.6 \times 10^7\text{ cm}^{-3}$). The evaporation rate increases by a factor of 10 for each $1.27\text{ kcal mol}^{-1}$ reduction of the cluster formation energy.

Extended Data Table 1 | Quantum chemical calculations of ELVOC cluster formation energies

Neutral clusters		Negative clusters		Positive clusters	
Cluster process	ΔG_{278K} (kcal/mol)	Cluster process	ΔG_{278K} (kcal/mol)	Cluster process	ΔG_{278K} (kcal/mol)
$E_1 + E_1$	-5.76	$E_1 + E_1^-$	-20.95	$E_1 + \text{NH}_4^+$	-22.46
$E_2 + E_1$	-2.15	$E_2 + E_1^-$	-19.90	$E_2 + \text{NH}_4^+$	-30.87
				$E_1 + E_1.\text{NH}_4^+$	-11.71
				$E_2 + E_1.\text{NH}_4^+$	-24.35
				$E_1 + E_2.\text{NH}_4^+$	-15.94
$E_1 + \text{H}_2\text{SO}_4$	-9.90	$E_1 + \text{HSO}_4^-$	-26.97		
$E_2 + \text{H}_2\text{SO}_4$	-12.04	$E_2 + \text{HSO}_4^-$	-30.30		
$E_1 + E_1.\text{H}_2\text{SO}_4$	+2.49	$E_1 + E_1.\text{HSO}_4^-$	-15.28		
$E_2 + E_1.\text{H}_2\text{SO}_4$	+3.13				
$E_1 + E_2.\text{H}_2\text{SO}_4$	-5.69				
		$E_1 + \text{NO}_3^-$	-25.99		
		$E_2 + \text{NO}_3^-$	-25.65		
		$E_1 + E_1.\text{NO}_3^-$	-10.09		

Formation Gibbs free energies at 278 K, ΔG_{278K} , for neutral, negatively charged and positively charged ELVOC clusters. The cluster processes indicate the incident E_1/E_2 vapour molecule + the target cluster. Quantum chemical calculations made at other temperatures (not shown) indicate that the binding energies strengthen by $-1.0 \text{ kcal mol}^{-1}$ per 20 K reduction in temperature. The uncertainty in the calculated energies is less than 2 kcal mol^{-1} . Our calculations indicate the following approximate order for different functional groups to contribute to the cluster binding energies involving HSO_4^- or H_2SO_4 (starting with the strongest): (i) carboxylic acids, $\text{R}-\text{C}(=\text{O})-\text{OH}$; (ii) hydroxyls, $\text{R}-\text{OH}$; (iii) hydroperoxy acids, $\text{R}-\text{C}(=\text{O})-\text{O}-\text{OH}$; (iv) hydroperoxides, $\text{R}-\text{O}-\text{OH}$; and (v) carbonyls, $\text{R}-(\text{R}')-\text{C}=\text{O}$. In the case of NH_4^+ , the main interacting group is carbonyl, independently of which other groups are attached to it; therefore ammonium will form stronger clusters with carboxylic acids, hydroperoxy acids or carbonyls than it will with hydroxyls or hydroperoxides.