Unveiling an Unexpected Branching Point in α-pinene Ozonolysis via Molecular Dynamics Guided Reaction Discovery

Huan Yang[1†#*], Umberto Raucci[2†*], Siddharth Iyer[3†], Galib Hasan[4], Thomas Golin Almeida[4], Shawon Barua[3], Anni Savolainen[3], Juha Kangasluoma[1], Matti Rissanen[3,4], Hanna Vehkamäki[1], Theo Kurtén[4*]

[1] Institute for Atmospheric and Earth System Research/Physics, University of Helsinki, FI-00014 Helsinki, Finland
[2] Atomistic Simulations, Italian Institute of Technology, 16156 Genova, Italy
[3] Aerosol Physics Laboratory, Tampere University, FI-33720 Tampere, Finland
[4] Department of Chemistry, University of Helsinki, FI-00014 Helsinki, Finland
[#] Current Address: Max Planck Institute for Chemistry, DE-55122 Mainz, Germany

Contributed equally (†)
Corresponding authors (*)

Abstract

Secondary organic aerosols (SOAs) significantly impact Earth's climate and human health. Although the oxidation of volatile organic compounds (VOCs) has been recognized as the major contributor to the atmospheric SOA budget, the mechanisms by which this process produces SOA-forming highly oxygenated organic molecules (HOMs) remain unclear. A major challenge is navigating the complex chemical landscape of these transformations, which traditional hypothesis-driven methods fail to thoroughly investigate. Here, we explored the oxidation of α-pinene, a critical atmospheric biogenic VOC, using a novel reaction discovery approach based on *ab initio* molecular dynamics and state-of-the-art enhanced sampling techniques. Our approach successfully identified all established reaction pathways of α-pinene ozonolysis, as well as discovered multiple novel species and pathways without relying on *a priori* chemical knowledge. In particular, we unveiled an unexpected branching point that leads to the rapid formation of alkoxy radicals, whose high and diverse reactivity help to explain hitherto unexplained oxidation pathways suggested by mass spectral peaks observed in α-pinene ozonolysis experiments. This branching point is likely prevalent across a variety of atmospheric VOCs and could be crucial in establishing the missing link to SOA-forming HOMs.

Organic aerosols are ubiquitous in the Earth's atmosphere and have significant impacts on human welfare[1-3]. On a regional level, they contribute to the formation of urban haze[4,5], which reduces visibility and endangers human health by increasing the risk of respiratory and cardiovascular diseases[6,7]. At the global scale, they affect the Earth's radiative balance through both direct and indirect mechanisms[3,8], with the former involving the scattering and absorption of solar radiation while the latter altering cloud properties. Model studies indicate that organic aerosol in general, and secondary organic aerosol (SOA) in particular[9-11], constitute a substantial fraction of the total number of aerosol particles in the atmosphere[3,12]. Unlike primary organic aerosol (POA) that are directly emitted in particulate form, SOAs form through the clustering of low-volatility compounds generated by the oxidation of gas-phase parent volatile organic compounds (VOCs), which in turn originate from both biogenic and anthropogenic sources such as vegetation or traffic. Some of these oxidation reactions can lead to highly oxygenated organic molecules (HOMs) with extremely low volatility, that can either partition into preexisting aerosols, thus increasing their size and mass, or even foster the formation of entirely new aerosol particles[13-18].

Traditionally, the oxidation of atmospheric parental VOCs was believed to proceed in a stepwise manner, with one oxidant (such as ozone and OH radical) attack adding at most 2 to 3 new oxygen-containing functional groups to the carbon backbone before forming closed-shell products. These products typically require further oxidation cycles to generate HOMs. This conventional perspective, however, strongly limits the potential yields of low-volatility HOMs, as the intermediate products are prone to being scavenged by partitioning into condensed-phase before undergoing another oxidation cycle. Recent research[15] has demonstrated that under suitable conditions, certain VOCs can undergo autoxidation, wherein a single oxidant attack leads (via sequential unimolecular isomerization and O_2 addition reactions) to products containing as many as 10 new oxygen atoms. These findings underscore the need for a more comprehensive understanding of parent VOC reactivities and their fates, especially the competition between autoxidation and sequential oxidation. Such insights are crucial for accurately assessing regional and global aerosol budgets and quantifying their consequential impacts on climate and health.

Except for polluted urban regions[19-21], biogenic VOCs have been identified as the main precursors for the global SOA budget, and total emissions of known biogenic SOA precursors are estimated to be up to 10 times larger than their anthropogenic counterpart[3]. Monoterpenes, a family of $C_{10}H_{16}$ compounds, comprise around 11% of global biogenic VOC emissions by mass[12]. They

are considered as a crucial class of precursors for the global SOA budget[12,22,23], primarily owing to their high reactivity (in particular, their tendency to undergo autoxidation) and the low volatility of the numerous oxidation products. α-Pinene is the most emitted monoterpene, accounting for approximately one third of the total emissions[24]. Its ozonolysis reaction (α-pinene + O_3) is one of the most efficient systems for SOA generation in the atmosphere[25,26]. Though significant efforts have been made in understanding this important reaction over the past decades[13,16,26-31], a conclusive characterization linking α-pinene ozonolysis to the SOA-forming lowest-volatility HOMs has not been reported[32].

The initial steps of α-pinene ozonolysis reaction, and the formation of closed-shell oxidation products with up to 4 to 5 oxygen atoms, are well documented and included, for example, in the widely used Master Chemical Mechanism[33] (MCM) for tropospheric degradation of VOCs. Further sequential oxidation of α-pinene is also reasonably well described by automated reaction mechanism generators such as the Generator of Explicit Chemistry and Kinetics of Organics in the Atmosphere[34,35] (GECKO-A). In contrast, the autoxidation pathways observed to lead to substantial yields of aerosol-forming HOMs[13,36] on a sub-second timescale are still poorly understood despite some promising recent advances[26]. An issue complicating the experimental analysis is that especially these autoxidation channels are only competitive in relatively clean conditions (low reactant concentrations) – at the high concentrations required for the application of spectroscopic techniques, for example, they are overwhelmed by termination reactions. Currently, mass spectrometry is the principal analytical technique in identifying and characterizing autoxidation pathways driving HOMs formation. However, without an added dimension of separation[37], mass spectrometric techniques cannot distinguish different molecules with the same mass. Computational methods complement mass spectrometry by identifying the actual molecular-level reaction steps behind the observed autoxidation process. However, thus far, only experimental observation-based hypothesis-driven computational approaches that rely heavily on chemical intuition have been used to identify autoxidation pathways. This limits the exploration of complex reaction networks characterized by multiple competitive reaction pathways.

In contrast to the hypothesis-driven strategy, several discovery-based approaches have recently been proposed to characterize the reaction space based on *ab initio* molecular dynamics (MD)[38-44]. Some of these methods employ high temperatures and pressures to induce chemical reactions[38,43,44], whereas others accelerate the sampling of reactive events by applying an external

bias potential to a set of collective variables (CVs) that encode the generally slow degrees of freedom of a system[39,40]. In this context, Raucci et al.[40] proposed a workflow for reaction discovery based on a CV derived from spectral graph theory[45] and the explore version of the on-the-fly probability enhanced sampling method[46] (OPES$_E$). In this approach, a molecule is represented as a graph whose vertices and edges are its atoms and chemical bonds, respectively. The maximum eigenvalue of the symmetric adjacency matrix associated with the graph is then used as CV in discovery simulations (Figure 1). This approach is suitable for blind sampling of reactive events because it is based on a generic CV that does not require any *a priori* knowledge about the reactivity of the system, allowing molecular dynamics simulations to freely explore new chemical pathways[40,47,48]. Furthermore, the dynamics of the system is perturbed in a controlled way by a barrier cutoff in OPES$_E$, which sets a limit on the maximum value of the bias, and grants control over the exploratory phase.

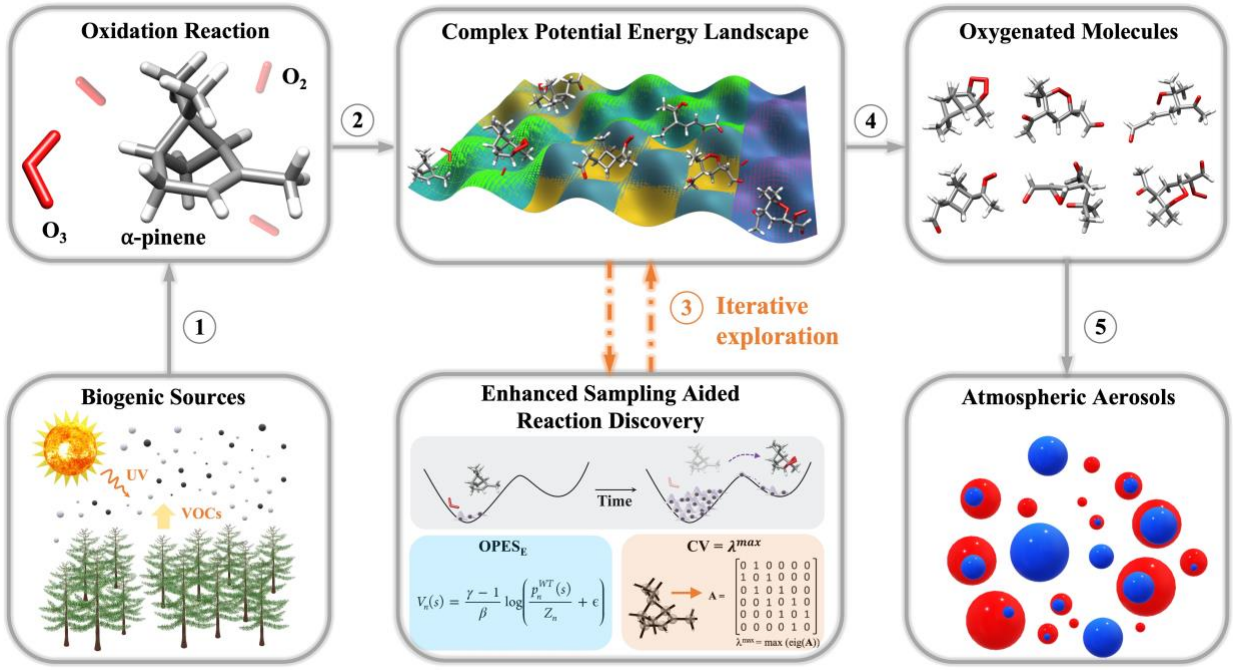

Figure 1. Schematic of the Reaction Discovery Approach. The intricate potential energy landscape arising from the degradation of atmospheric VOCs is systematically explored through the enhanced sampling aided approach, facilitating the enumeration of potential reaction pathways for subsequent high-level quantum chemical calculations.

In this study, we unravel the complexity of α-pinene ozonolysis using this enhanced sampling aided reaction discovery approach. In a three-step strategy, we first performed extensive

MD guided reaction discovery simulations, and then characterized the energetics of the newly discovered pathways with high-level *ab initio* electronic structure calculations (up to CCSD(T) level). Finally, a kinetic master equation approach was set up to understand the significance of the newly observed reaction routes.

Without relying on human heuristic, we were able to uncover all the known reaction pathways of the α-pinene ozonolysis reaction, including the ring-opening channel and the OH roaming channels reported very recently by Iyer et al.[26] and Klippenstein and Elliott[49], respectively, as well as discover several new species and reaction pathways that are of high atmospheric relevance. In particular, our approach revealed the importance of a new reaction class in α-pinene ozonolysis: the unimolecular rearrangement of alkyl radicals containing endoperoxide functionalities, resulting in the formation of alkoxy radicals. This reaction type was investigated in aromatic oxidation but found to be negligible for benzene[50]. In contrast, within the α-pinene ozonolysis system, this reaction type becomes activated due to the excess energy retained from preceding reaction steps. In addition, by exploiting the analogy between this reaction and a previously reported epoxidation reaction[51], we have identified both anthropogenic and biogenic model systems in which the endoperoxy-alkyl rearrangement is competitive with O_2 addition even as a thermal reaction, i.e., without the chemical activation driven by the excess energy. A key feature of this reaction is that it unimolecularly converts alkyl radicals into alkoxy radicals, and thus – due to the versatility of alkoxy radical atmospheric chemistry – in effect opens new and unanticipated branching points in atmospheric oxidation sequences. Considering the ubiquity of endoperoxide functionalities in aromatic compounds and terpene-derived intermediates, such branching points will offer crucial evidence in explaining the extensive array of as-yet unexplained molecular compositions observed in mass spectrometric autoxidation experiments, and in delineating their role in atmospheric aerosol formation.

Results

Thorough exploration of known α-Pinene ozonolysis pathways via MD guided reaction discovery

The initial steps of the α-pinene ozonolysis reaction follows the general mechanism of cyclic alkene ozonolysis. The ozone molecule adds to the double bond, forming a primary ozonide,

which rapidly decomposes into a carbonyl-substituted Criegee Intermediate (CI; carbonyl oxide) (Figure 2). Due to the asymmetry of the addition site, as well as the high barrier for rotation around the O-O bond in the CI, there are four distinct CI isomers. According to MCM[33], the branching ratios for their formation are reasonably similar[33] (between 0.33 and 0.2). Rapid 1,4 hydrogen shift channels are available to three of the four CI isomers, leading to vinyl hydroperoxide (VHP). These in turn lose an OH radical to form "vinoxy" radicals (resonance-stabilized species with both alkyl and alkoxy radical character), which add O_2 to form the first-generation peroxy radicals (RO_2) (Figure 2). The remaining CI isomer cannot undergo H-shifts, and thus either reacts bimolecularly (for example with water vapor) or isomerizes into a dioxirane (Figure S1a) which then rearranges or decomposes further to a variety of closed-shell products including carboxylic acids. This "textbook" reaction sequence is unable to explain the experimentally observed autoxidation in the α-pinene ozonolysis system, as all of the isomerization reactions of the first-generation RO_2 are far too slow to compete with bimolecular termination even in the cleanest conditions.[30] Based on theoretical calculations by Kurtén et al.[30], these isomerization reactions can at most add one more O_2 to the carbon chain under atmospheric conditions. In 2021, Iyer et al.[26] partially resolved this problem by introducing an alternative reaction, where one of the vinoxy radicals undergoes a chemically activated ring-opening, followed by O_2 addition to make a new RO_2 radical capable of undergoing rapid unimolecular isomerization reactions. Experiments on selectively deuterated α-pinene isomers[32] have later confirmed that this reaction channel is indeed competitive – but also demonstrate the existence of yet more still undiscovered reaction channels. Specifically, while the reaction route forming closed-shell products with 7 or more O atoms in Iyer et al[26], such route inevitably involve H-shifts of the aldehydic H atom attached to the carbon that was originally the secondary C atom in the double bond (carbon 3 in Meder et al.[32]). But the deuteration experiments indicate that almost half of these products have not undergone such a reaction. This highlights the existence of unaccounted processes essential for explaining the α-pinene ozonolysis experiments.

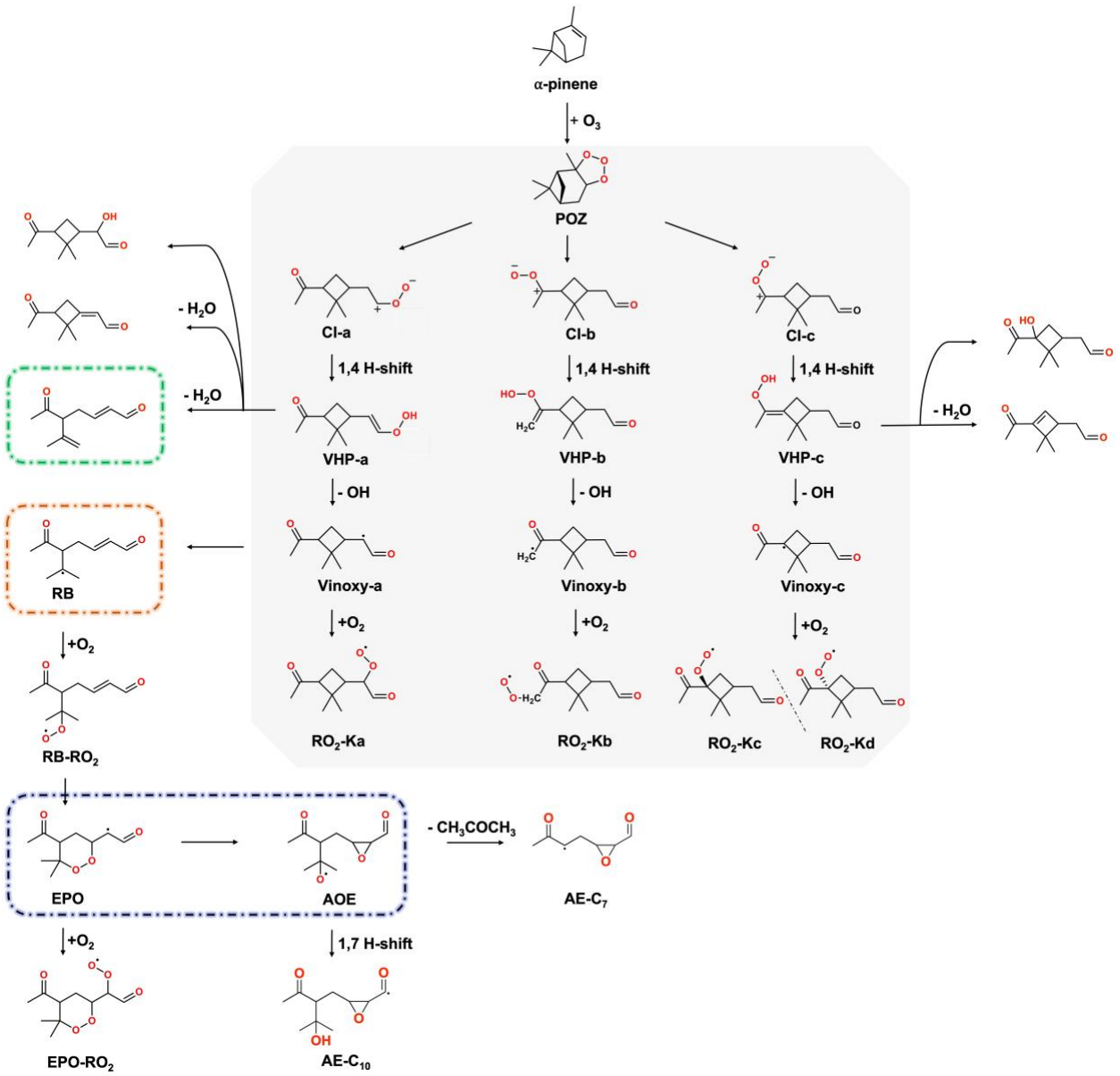

Figure 2. α-pinene ozonolysis mechanisms from the MD driven reaction discovery simulations. The grey region represents the conventional perception of the α-pinene ozonolysis mechanisms. The ring-breaking mechanism proposed by Iyer et al.[26] is marked orange. The OH roaming mechanism leading to a closed-shell product with two double bonds found in this work is marked green. The endoperoxy-alkyl radical rearrangement mechanism found in this work is marked blue.

The "textbook" α-pinene ozonolysis mechanism[33] described above is reproduced in full by the MD reaction discovery procedure, illustrated in gray in Figure 2. Moreover, the reaction discovery engine independently discovered the ring-breaking mechanism (illustrated in orange in

Figure 2) and its main subsequent steps proposed by Iyer et al[26]. In addition to these already-known reaction channels, the sampling procedure produced a vast number of additional reaction mechanisms following the ring-breaking of Vinoxy-c, summarized in Figure S1b. While these reaction channels are also likely to be real (in the sense of being actual pathways on the potential energy surface of the α-pinene + O_3 + nO_2 system), the vast majority of them are not competitive with bimolecular sinks in the atmosphere, which typically constrain the lifetimes of alkyl radicals to less than 10^{-6} s (due to the reaction with O_2) and peroxy radicals to less than 0.01 s to 100 s (due to reactions with either NO_x or HO_2, depending on the conditions). The lifetimes and fates of alkoxy radicals are more structure-dependent[52], but inevitably less than 10^{-3} s.

Notably, the reaction discovery engine also independently identified the OH roaming pathway, which was recently proposed by Klippenstein and Elliott based on systematic theoretical calculations[49]. The OH roaming was thought to be responsible for the production of hydroxycarbonyl products during the dissociation of Criegee intermediates, as observed in the pioneering experimental work[53] and several recent works[54,55] by Lester and co-workers. In the context of α-pinene ozonolysis, Klippenstein and Elliott's theoretical calculations[49] indicate that the OH roaming channels are kinetically significant, with a branching ratio of around 20% at room temperature for stabilized VHPs, and this ratio can further increase at the lower temperatures typical of the troposphere. However, the overall impact of the OH roaming channels is constrained by the branching of stabilized VHPs. The initial step of OH roaming produces addition or abstraction products with intact cyclobutane rings, which could possess energy exceeding the dissociation threshold of the cyclobutane ring and then lead to ring-breaking products[49]. We found a new ring-breaking product with two carbon-carbon double bonds, highlighted in green in Figure 2. This product forms when the dissociated OH abstracts a hydrogen atom from one of the two spectator methyl groups, leading to cyclobutane ring cleavage. Such unsaturated products have a high potential for subsequent oxidation reactions and could thereby contribute to the formation of HOMs. Their oxidation paths and related products merit further in-depth studies in the future.

New branching point in α-pinene ozonolysis: endoperoxy-alkyl radical rearrangement

One recurring structural motif in the novel channels discovered by the reaction discovery engine was the high reactivity of alkyl radicals with endoperoxide (cyclic peroxide) substituents. These were observed to undergo a variety of different isomerization reactions, typically resulting

in the conversion of an alkyl radical with an endoperoxide group into an alkoxy radical with an epoxide or ether group. Due to lifetime constraint discussed above, most of these reaction channels are unlikely to be competitive in the atmosphere. Similar rearrangement mechanisms have been reported for other organic compounds. For instance, Zador et al.[56] reported this mechanism in a selected intermediate from the limonene + OH system. Additionally, calculations by Xu et al.[50] suggested that this rearrangement mechanism is excessively slow in the oxidation of benzene. Nevertheless, for the radical depicted in blue in Figure 2, this reaction channel may have significant yields due to large amount of excess energy retained from preceding reaction steps. Indeed, this novel reaction channel occurs downstream of the ring-breaking mechanism proposed by Iyer et al.[26], and involves two new isomerization steps for the RB-RO_2 species. The first step, with a barrier of 16.3 kcal/mol at the CCSD(T) level, yields an alkyl radical with an endoperoxide group (abbreviated EPO in Figure 2). The second step, with a CCSD(T) barrier of 17.9 kcal/mol, then rearranges the EPO into an alkoxy radical with an epoxide group (abbreviated AOE).

To quantify the impact of excess energy in overcoming the above rearrangement channel barriers, we conducted comprehensive kinetic master equation simulations covering the entire potential energy surface of α-pinene ozonolysis along the targeted Criegee pathway (CI-a), depicted in the upper panel of Figure 3. Within these simulations, the dissipation of excess energy was addressed using a carefully evaluated molecular heat transfer model (Supplementary Materials section 2). Our simulations unveiled that over 75% of RB species was formed within 2 ns after the initial ozone attack on α-pinene (Figure S4), suggesting that RB species should retain a significant portion of the 80 kcal/mol excess energy acquired in preceding reactions. Traditionally, it has been assumed that such excess energy dissipates during the next O_2 addition step (given its relatively lengthy timescale of approximately 100 ns), thereby exerting minimal influence on subsequent reaction steps. However, this assumption is challenged by our current simulation results: even though excess energy may dissipate after the completion of O_2 addition, the early-formed addition product can retain sufficient energy to surmount the transition state (TS_5 and TS_6) barriers, leading to the substantial formation of AOE within nanoseconds. The yield of this endoperoxy-alkyl radical rearrangement channel exhibits a high sensitivity to the O_2 addition rate. Here, we used a model transition state with variable energies (while the O_2 concentration was fixed at 5×10^{18} cm^{-3}) in simulations to control the effective pseudo-unimolecular O_2 addition rate. We found that, at an effective pseudo-unimolecular O_2 addition rate of 1.9×10^{7} s^{-1}, the yield is approximately 9.0%,

while a rate of 9×10^7 s^{-1} elevates the yield to 31.5%, as is depicted in Figure 3. Assuming roughly equal yields for the four Criegee-forming channels of α-pinene ozonolysis, the 31.5% yield from the studied single channel translates into an overall yield of approximately 7.9% for the complete α-pinene ozonolysis reaction (this percentage roughly equals to the full HOMs yield from α-pinene ozonolysis). It is noteworthy that, in contrast to previous results[30], the excess energy may also accelerate the H-shift reaction of certain first-generation RO_2 species derived from α-pinene, enabling the nanosecond-scale formation of some H-shift products (Supplementary Materials section 4).

Moreover, the reactant alkyl radical EPO is resonance-stabilized, indicating potential reversibility of the competing O_2 addition reaction, and a further enhancement of the endoperoxy-alkyl radical rearrangement channel. However, additional kinetic master equation simulations suggest that the impact of this back-reaction becomes significant only when the lifetime of the O_2 addition product (denoted EPO-RO_2 in Figure 2) is unrealistically long, exceeding 10^5 s (Supplementary Materials section 5). Lowering the rearrangement reaction barrier to approximately 14 kcal/mol is necessary for the back reaction to exert influence under realistic RO_2 lifetimes of 1 s to 100 s. The dependence of the endoperoxy-alkyl radical rearrangement channel on the O_2 addition product lifetime provides a typical example of the competition between kinetic and thermodynamic control in O_2 addition to resonance-stabilized alkyl radicals, as discussed for example by the work of Wennberg and co-authors[57] in the context of isoprene oxidation. While not applicable to the α-pinene case, the O_2 addition back-reaction may play a role in other atmospherically relevant systems discussed in the subsequent section, for which the rearrangement reaction barriers are in the range of 11 kcal/mol to 14 kcal/mol.

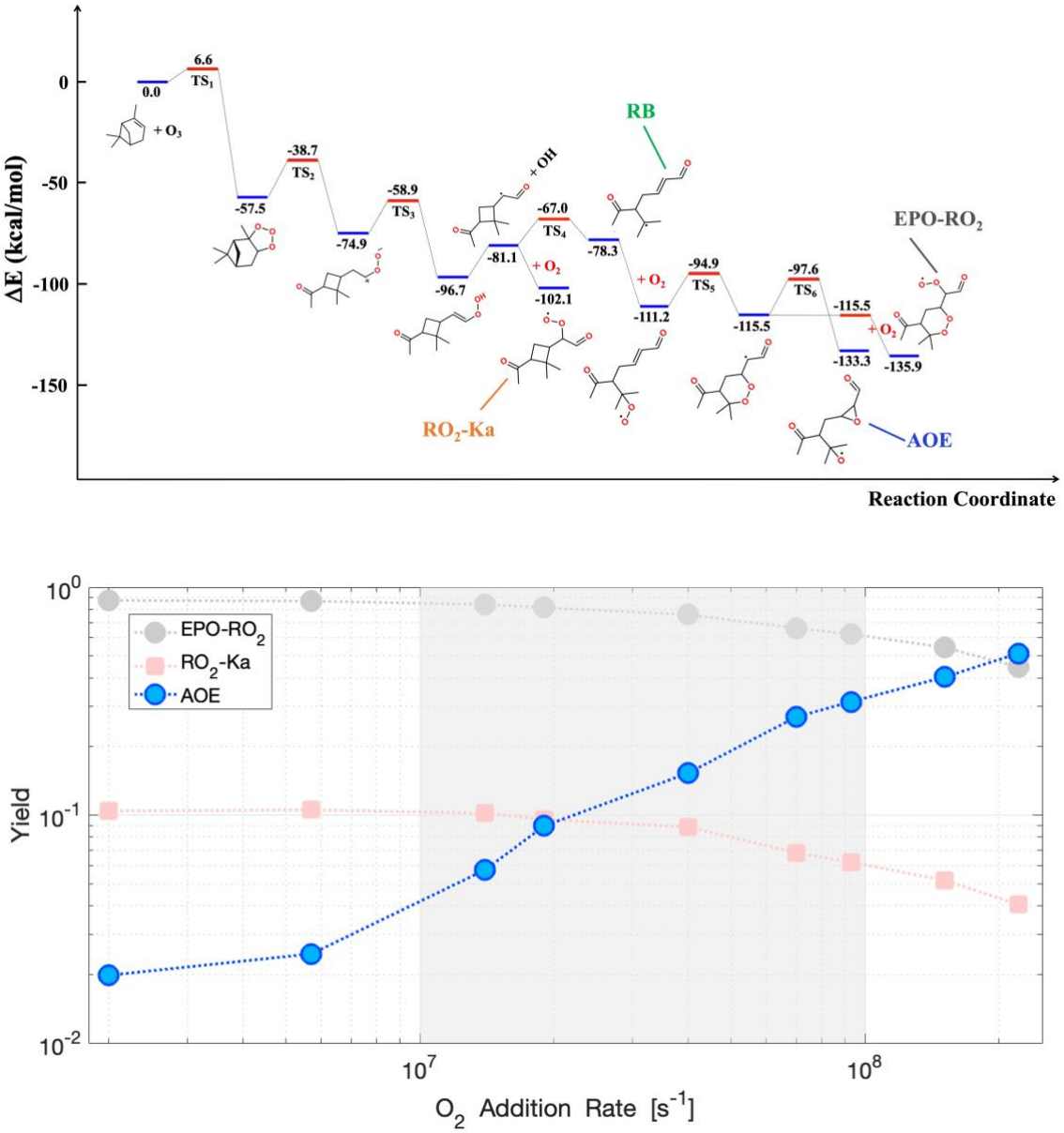

Figure 3. Upper panel: potential energy surface of α-pinene ozonolysis for the CI-a isomeric pathways. Numerical values are zero-point corrected energies as explained in the Methods section. Lower panel: yield of the endoperoxy-alkyl radical rearrangement reaction (AOE) following the CI-a isomeric pathway compared to its competing channels.

The substantial yield observed within the plausible pseudo-unimolecular O_2 addition rate range[51,58,59] of 10^7 s^{-1} to 10^8 s^{-1} suggests that the newly identified endoperoxy-alkyl radical rearrangement mechanism could be a major channel in α-pinene ozonolysis. The existence of this rearrangement channel is supported by recent flow reactor experiments[60] which show that the yield

of 8-oxygen containing acyl peroxy radical is likely lower than estimated by Iyer et al.[26]. This indicates an alternative fate of one or more of the preceding intermediates, such as the rearrangement of the endoperoxy-alkyl radical reported here. We note that in contrast to the peroxide product from the O_2 addition, the alkoxy radical product formed from the rearrangement reaction is likely to have a large number of rapid reaction channels available that could help explain the hitherto unexplained oxidation pathways suggested by the mass spectral peaks.

Potential routes to Highly Oxygenated Molecules

Potential follow-up reaction channels for the endoperoxy-alkyl radical rearrangement product AOE were explored (Supplementary Materials section 9). While the scission reaction leading to acetone and a C_7 product (AE-C_7 in Figure 4a) is the fastest (this reaction was also observed more frequently than others in the MD reaction discovery simulations), several H-shift channels are likely to have non-negligible yields. In particular, kinetic master equation simulations using the DFT barriers in Table S4 suggest that at 300 K the 1,7 H-shift of AOE (producing AE-C_{10}) has a rate roughly one tenth that of its scission reaction (producing AE-C_7), converting to a branching ratio of about 0.1 v.s. 0.9 for the two paths. O_2 addition products (RO_2) following AE-C_7 and AE-C_{10} contain multiple active sites for H-shifts (e.g., H atoms on the carbonyl carbon of the aldehyde and on the carbons of the epoxide ring) and hence have a strong potential for further autoxidation. According to the recently published structure activity relationship by Vereecken and Noziere[61], H-shifts on the aldehyde exhibit rates exceeding 1 s^{-1}. We are unaware of any specific structure activity relationship for H-shifts of RO_2 containing an epoxide group. As a rough estimate, we can assume that H atoms on cyclic ether (epoxide) carbons have similar reactivity as H atoms on carbons with an acyclic ether substituent. For these, the structure activity relationship[62] predicts H-shift rates up to 0.1 s^{-1}. Our DFT calculations suggest that for AE-C_7-RO_2, the H-shifts from the aldehyde and epoxide carbons have barriers of around 22 - 25 kcal/mol, corresponding to rates exceeding 0.01 s^{-1} at the low barrier end. These rates align with the assumption that H-shifts from cyclic (epoxide) and acyclic ether carbons have similar rates. However, our calculations predict somewhat lower rates than the structure activity relationship[61] for H-shifts from the aldehyde carbon. This is likely due to the steric hindrance caused by the epoxide ring on the carbon chain in the current case. Nevertheless, these rates suggest that H-shift reactions following AE-C_7 and AE-C_{10} could be competitive with their bimolecular sinks. Figure 4a shows potential routes

for autoxidation leading to HOMs with up to 10 or 11 oxygen atoms. Note that these routes are only illustrative, as the order of H-shifts may vary due to their similar rates. Additionally, some of the H-shift products from the epoxide (e.g., the 1,4 H-shift of AE-C_{10}-RO_2) will likely not retain the epoxide ring intact, but instead form a carbonyl with O_2 adding to the adjacent carbon. Many of these routes likely contribute to the mass spectrometry peaks for C_7 and C_{10} HOMs observed in our flow reactor experiments (Figure 4b and 4c).

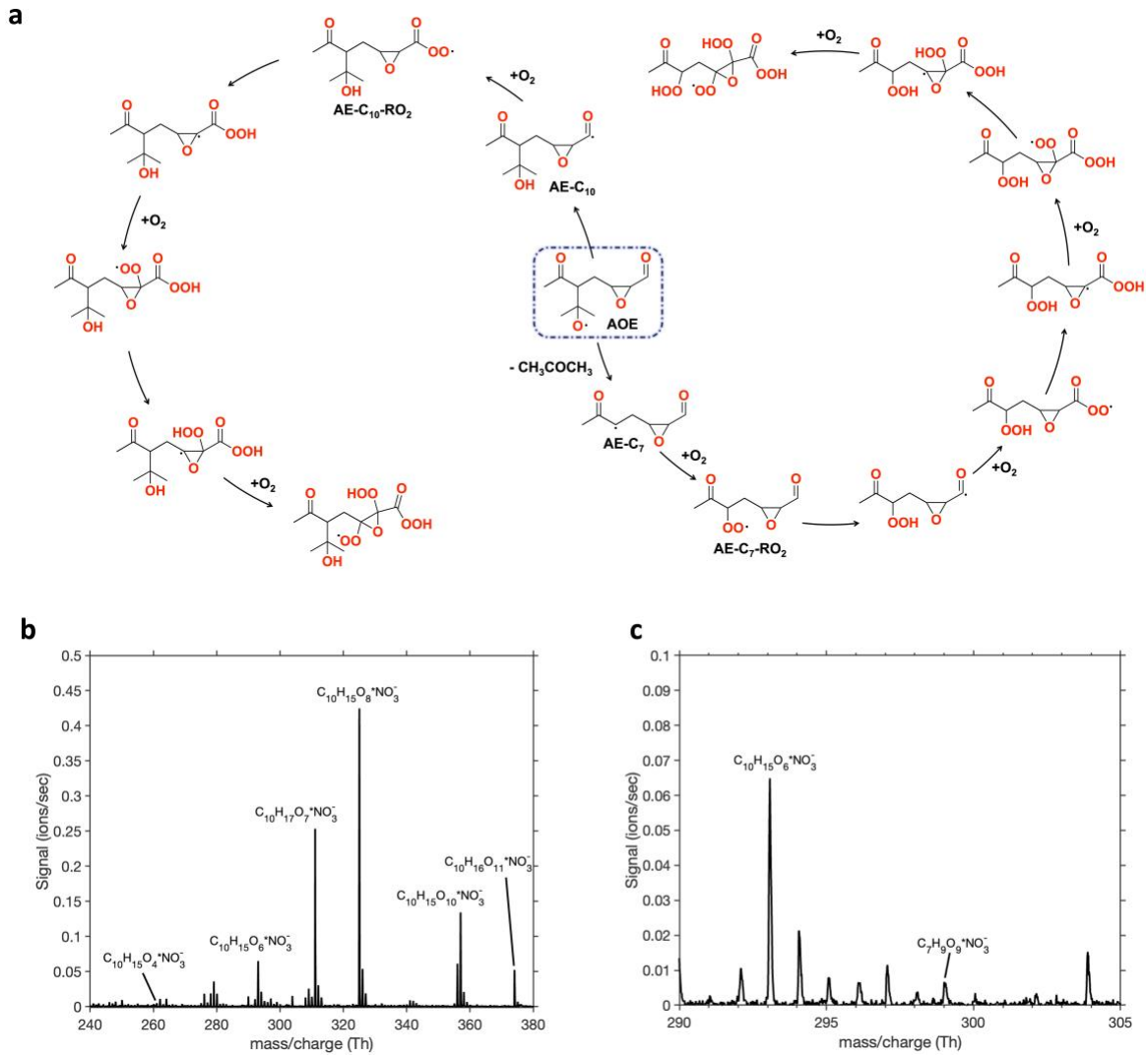

Figure 4. (a) Potential autoxidation routes for AE-C_7 and AE-C_{10}; reactions after AE-C_7-RO_2 and AE-C_{10}-RO_2 are based on structure activity relationship and chemical intuition. (b) Mass Spectrometry for C_{10} HOMs observed at the reaction time of 600 ms. (c) Mass Spectrometry for C_7 HOM observed at the reaction time of 600 ms.

Endoperoxy-alkyl radical rearrangement for other atmospheric compounds

The endoperoxy-alkyl radical rearrangement mechanism discovered here can be thought of as an extension of the epoxide-forming channel investigated early by Møller et al.[51] Direct experimental observation of that reaction was also achieved in the notable work of Klippenstein, Lester, and their co-workers[63]. In that reaction, an alkyl radical adjacent to a hydroperoxy group reacts to form a closed-shell epoxide and a OH radical. Møller et al.[51] found that the thermal reaction rate is competitive with O_2 addition when the alkyl radical carbon has an OH substituent, which can form a H-bond to an acceptor substituent on or adjacent to the other C atom in the nascent epoxide ring. We carried out test calculations on a variety of systems (replacing the OOH group with a cyclic peroxide) and found that the presence of the OH substituent on the alkyl radical carbon can indeed significantly reduce the barrier of the newly discovered endoperoxy-alkyl radical rearrangement channel. The corresponding reaction barriers for most of these tailored molecules are approximately 11 kcal/mol to 14 kcal/mol at the CCSD(T) level (Supplementary Materials section 7), in contrast to the 17.9 kcal/mol at the CCSD(T) level observed in the case of α-pinene. While even these reduced barrier heights still do not correspond to competitive reaction rates against O_2 addition under thermalized conditions, simulations on model systems (Supplementary Materials section 6) indicate that the reactions could have non-negligible yields when considering a modest and plausible amount of excess energy (e.g., 20 kcal/mol) from preceding steps. Despite not identifying such an OH substituent on the alkyl radical carbon in the known or sampled α-pinene ozonolysis system, this structural motif is likely present in the atmospheric oxidation of a diverse array of anthropogenic and biogenic compounds.

We identified three representative examples with endoperoxy-alkyl radical rearrangement reaction rates potentially competitive against O_2 addition: the naturally emitted monoterpene α-ocimene (Figure 5), the industrial chemical 1,4 hexadiene (Figure S11a), and the terpenoid alcohol geraniol, which is both naturally emitted and used in fragrances (Figure S11b). For the 1,4 hexadiene + OH and geraniol + OH systems, the corresponding endoperoxy-alkyl radical rearrangement reaction barriers are 13.4 kcal/mol and 11 kcal/mol, respectively. Simulations on model systems suggest that such barrier heights could result in about 10% and 40% yields, respectively, when factoring in 20 kcal/mol of excess energy from preceding reaction steps. Considering the potential reversibility of O_2 addition, this yield could likely be increased further.

For the α-ocimene + OH system, the endoperoxy-alkyl radical rearrangement reaction barrier is only 3.1 kcal/mol at the CCSD(T) level, indicating a 100% yield even under thermalized conditions without any excess energy. This typical example in the α-ocimene + OH system highlights the importance of the new mechanism. Conventionally, hydroxyalkyl radicals are expected to terminate the oxidation propagation by releasing the H atom to an O_2, resulting in the formation of closed-shell carbonyl co-products. However, the new mechanism suggests that these intermediates do not necessarily conclude the oxidation propagation or the formation of HOMs.

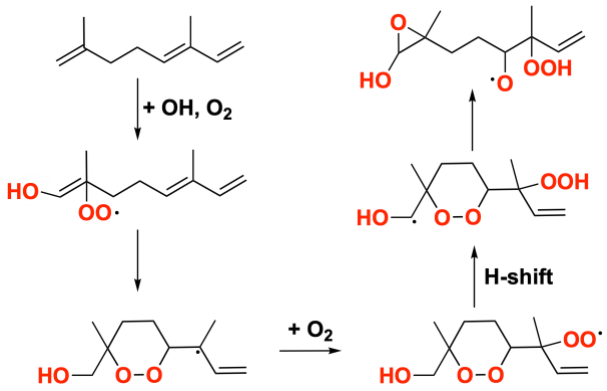

Figure 5. The endoperoxy-alkyl radical rearrangement reaction channel within the α-ocimene ozonolysis system. The barrier for this endoperoxy-alkyl radical rearrangement reaction is 3.1 kcal/mol at the CCSD(T) level, indicating a much faster reaction rate and 100% yield against its competing O_2 addition reaction.

Discussion

These selected examples demonstrate that the endoperoxy-alkyl radical rearrangement mechanism presented here is likely to be relatively common in atmospheric autoxidation processes. OH groups, alkyl radical sites, endoperoxides and H-bond acceptor groups are ubiquitous functionalities – the specific structural combination required to make the proposed reaction pathway competitive is not implausibly restrictive. The main atmospheric implication of the reaction is that it opens new and unexpected branching points in oxidation mechanisms: while alkyl radicals, with a few exceptions, have traditionally been assumed to exclusively react with O_2, alkoxy radicals are arguably the most versatile species in atmospheric chemistry, and often have multiple competing reaction channels available. A recurring theme of both smog chamber and field studies of atmospheric oxidation has been the surprisingly large numbers of different

progressively oxygenated product compounds (e.g., in terms of elemental compositions in mass spectra). The unexpected branching points revealed by our new mechanism helps explain this diversity.

Acknowledgments

We thank Prof. Michele Parrinello for reading a preliminary version of the paper and for his encouraging remarks and support. We also thank Valerio Rizzi, Axel Tosello Gardini, and Enrico Trizio for helpful discussions. This work was supported by The Jane and Aatos Erkko Foundation (JAES) and the Reserch Council of Finland Center of Excellence VILMA under Grant Nos. 346369 and 346368. M.R. acknowledges the funding support from the European Research Council under the European Union's Horizon 2020 research and innovation programme under Grant No. 101002728. We thank the CSC IT center for science for computing resources.

Methods

Enhanced sampling aided reaction discovery

Our approach for conducting reaction discovery simulations combines a collective variable (CV) derived from spectral graph theory[40] and the explore variant of the On-the-fly Probability Enhanced Sampling ($OPES_E$) method.[46] The workflow is thoroughly explained in a previous publication[40] and only briefly outlined here. The molecule is represented as a graph where vertices and edges represent atoms and chemical bonds, respectively. The adjacency matrix **A** is associated to the graph, and its elements a_{ij} indicates whether atoms i and j are connected by a chemical bond:

$$a_{ij} = \frac{1-\left(\frac{r_{ij}}{\sigma_{ij}}\right)^{n}}{1-\left(\frac{r_{ij}}{\sigma_{ij}}\right)^{m}} \quad (1)$$

where σ_{ij} represent the typical bond lengths between atoms of types i and j. We use the largest eigenvalue of the adjacency matrix **A** (denoted as λ^{max}) as collective variable in the enhanced sampling simulations. λ^{max} is real, positive, and non-degenerate, and its choice as a CV guarantees translational, rotational and permutational invariance while preserving the chemical information encoded in the molecular graph (λ^{max} grows with the number of bonds and its value lies between the average and maximum coordination number[45]).

The fluctuations of λ^{max} are enhanced using the $OPES_E$ method, which builds the bias by estimating the probability distribution of the CV on-the-fly. In particular, the bias is built as[46]:

$$V_n(s) = \frac{\gamma-1}{\beta}\log\left(\frac{p_n^{WT}(s)}{Z_n} + \varepsilon\right) \quad (2)$$

where Z_n is a normalization factor, $\gamma > 1$ is the bias factor, and $p^{WT}(s)$ represents the well-tempered distribution. An important feature of $OPES_E$ is the presence of a regularization term ε ($\varepsilon = e^{-\beta\Delta E/(1-1/\gamma)}$) which is related to the maximum value of the bias that can be deposited. This allows limiting the extent of exploration by preventing high energy transition states from being visited, and thus, represents a key parameter to control the discovery phase.

The discovery engine is complemented by a reaction analysis module that enables detection of new chemical species. We use the Open Babel[64] program to identify chemical species in every trajectory snapshot, converting the Cartesian coordinates of each frame into their corresponding

SMILES representations. The adjacency matrix **A** includes all the atoms present in the system. **A** elements (a_{ij}) are computed using PLUMED[65], with the σ_{ij} parameters of Eq. 1 being equal to 1.6 Å for O-H, 1.9 Å for O-O, C-C, and C-O, and 1.2 Å for C-H and H-H. The exponents of the switching functions are chosen as n = 6 and m = 10 to enforce a smooth behavior over a wide range of distances. In the case of the VHP species, defining the graph with all the atoms in the system predominantly resulted in sampling various conformers. Therefore, we conducted simulations using a CV more tailored on a specific functional group that undergoes the OH decomposition (i.e., the -COOH group in the VHP), observing several reactive pathways involving the formation of OH radicals.

We started the discovery phase from α-pinene and ozone. We performed multiple molecular dynamics (0.5 ns each) setting the barrier value in OPES explore at 100, 200, 300, 400, and 500 kJ/mol. Whenever a new species was detected, we initiated new simulations from that structure, again with multiple values of the ε parameter in $OPES_E$. Each new simulation was carried out both in the presence and absence of O_2. Each discovered species was first optimized at DFT level (B3LYP/6-31+G(d)) to assess its stability. Then, the most relevant pathways were characterized by computing transition states (TS) at a high level of theory (CCSD(T)), as explained in the following section. Molecular dynamics simulations were performed using the CP2K 8.11[66] software, combined with a development version of PLUMED 2.8.2.[65] All simulations of the discovery phase were conducted at the PM6[67] level of theory, with an integration time step of 0.5 fs. We sampled the NVT ensemble using the velocity rescaling thermostat,[68] maintaining a temperature of 300 K and a time constant of 100 fs.

Electronic structure calculations

A thorough exploration of the reactants, intermediates, transition states, and products was conducted by systematically sampling their conformers using the Spartan'18 (Wavefunction, Inc) program. The conformer sampling procedure was carried out using the MMFF method[69]. All conformers were first optimized at the B3LYP/6-31+G(d) level of theory[70-73], and those within 2 kcal/mol in relative electronic energies of the lowest energy conformer were chosen to be optimized at the ωB97X-D/aug-cc-pVTZ level of theory[74-76]. Finally, only the lowest energy conformer was chosen for the final single-point electronic energy calculation at the ROHF-ROCCSD(T)-F12a/VDZ-F12 level[77-81] to correct the final energies. These calculations were

performed using the Molpro 2022.2.2 program.[82] To obtain transition state (TS) structures, initial guesses were generated in Spartan'18 and subsequently optimized at the B3LYP/6-31+G(d) level of theory using Gaussian 09, with relevant TS bond distances constrained. The selection of the B3LYP functional for the initial TS calculations was based on its established capability to determine TS structures. Unconstrained TS optimization at the same level of theory was then performed immediately after the constrained optimization of the initial guess TS structures to precisely locate the TS structures. Once the TS structures were identified, they underwent conformer sampling in Spartan with relevant bonds being constrained. The resulting TS conformers were optimized once again under constraint and subsequently subjected to unconstrained TS optimization at the B3LYP/6-31+G(d) level in Gaussian 16. Finally, conformers with relative electronic energies within 2 kcal/mol were optimized at the higher ωB97X-D/aug-cc-pVTZ level of theory. The lowest energy conformer was selected for the final single-point electronic energy calculation at the ROHF-ROCCSD(T)-F12a/VDZ-F12 level using the Molpro program[82].

Kinetic master equation calculations

The RRKM simulations were carried out using the Master equation solver for multi-energy well reactions (MESMER) program.[83] The potential energy surface and the methods used until the formation of the ring broken RO_2 are identical to those used in our previous work[26]. Briefly, the SimpleRRKM method in MESMER with Eckart tunneling was used to treat the initial association reaction. Similar treatment was applied to the other intermediate complexes separated by transition states. Approximations of the nascent energy allocation to both vinoxy and OH products resulting from VHP decomposition were tackled employing the pseudo-isomerization methodology[84]. The parameter *m* value, which controls the fraction of the density of states of the vinoxy radical to spread the excess energy over, was set to 0.5. The O_2 addition step of the ring-broken product RB was treated using the SimpleRRKM method, employing a model transition state with variable energies (while O_2 concentration was fixed at 5×10^{18} cm^{-3}) to modulate the effective pseudo-unimolecular addition rate within the range of about 10^7 s^{-1} to 10^8 s^{-1}, a plausible rate range for O_2 addition to vinoxy/alkyl radicals under atmospheric conditions[51,58,59] (T = 298.15 K and P = 1 atm). The O_2 addition step for all other vinoxy/alkyl radicals was managed using the "Simple Bimolecular Sink" method, wherein the rate is governed by an adjustable bimolecular loss rate

coefficient and an O_2 "excess reactant" concentration of 5×10^{18} cm^{-3}. The dissipation of the excess energy through interaction with bath gases was regulated by the ΔE_{down} parameter in MESMER and was set to be 175 cm^{-1}. This selection was guided by a molecular heat dissipation model discussed in Section 2 of the Supplementary Materials. All intermediates were designated as "modeled" throughout the simulations, with corresponding Lennard–Jones potentials assigned as sigma = 6.5 Å and epsilon = 600 K. The simulation employed a grain size of 120 cm^{-1}, with the energy spanned by the grains of 60 k_BT. The initial ozone concentration was set identical to our prior work, at 10^{18} molecules cm^{-3}, to minimize the overall simulation time. Previous test calculations[26] run with the more ambiently relevant ozone concentration (10^{12} molecules cm^{-3}) have shown that the initial ozone concentration did not influence results. This rationale aligns with the observation that, beyond the initial association reaction, subsequent ozonolysis steps exhibit independence from the initial ozone concentration.

Mass spectrometry experiments

A time-of-flight chemical ionization mass spectrometer with nitrate charger ions (NO_3-CIMS) was used to detect the oxidation products of α-pinene ozonolysis at 600 ms reaction time. The experiments were carried out in a quartz tube flow reactor. Liquid α-pinene (Sigma-Aldrich, purity 91%) was bubbled from a reservoir using a pure N_2 flow. This was allowed to interact with ozone produced by flowing synthetic air through an ozone generator equipped with a 184.9 nm (Hg PenRay) lamp. The reaction time between α-pinene and ozone was controlled in two ways, 1) by introducing the ozone into the flow reactor through an injector tube and consequently limiting the α-pinene and ozone interaction time, and 2) by increasing the inlet flow through the reactor to the mass spectrometer. High inlet flow rates of 20 – 30 liters per minute to achieve sub-second reaction times were made possible by the use of the Multi-scheme chemical ionization inlet[85] (MION). The experiments were conducted under atmospheric conditions.

References

1 Jimenez, J. L. *et al.* Evolution of organic aerosols in the atmosphere. *science* **326**, 1525-1529 (2009).

2 Pöschl, U. Atmospheric aerosols: composition, transformation, climate and health effects. *Angewandte Chemie International Edition* **44**, 7520-7540 (2005).
3 Kanakidou, M. *et al.* Organic aerosol and global climate modelling: a review. *Atmospheric Chemistry and Physics* **5**, 1053-1123 (2005).
4 Guo, S. *et al.* Elucidating severe urban haze formation in China. *Proceedings of the National Academy of Sciences* **111**, 17373-17378 (2014).
5 An, Z. *et al.* Severe haze in northern China: A synergy of anthropogenic emissions and atmospheric processes. *Proceedings of the National Academy of Sciences* **116**, 8657-8666 (2019).
6 Al-Kindi, S. G., Brook, R. D., Biswal, S. & Rajagopalan, S. Environmental determinants of cardiovascular disease: lessons learned from air pollution. *Nature Reviews Cardiology* **17**, 656-672 (2020).
7 Pope III, C. A. & Dockery, D. W. Health effects of fine particulate air pollution: lines that connect. *Journal of the air & waste management association* **56**, 709-742 (2006).
8 Kiehl, J. Aerosols, climate, and the hydrological cycle. *science* **294**, 2119-2124 (2001).
9 Hallquist, M. *et al.* The formation, properties and impact of secondary organic aerosol: current and emerging issues. *Atmospheric chemistry and physics* **9**, 5155-5236 (2009).
10 Shrivastava, M. *et al.* Recent advances in understanding secondary organic aerosol: Implications for global climate forcing. *Reviews of Geophysics* **55**, 509-559 (2017).
11 Ziemann, P. J. & Atkinson, R. Kinetics, products, and mechanisms of secondary organic aerosol formation. *Chemical Society Reviews* **41**, 6582-6605 (2012).
12 Guenther, A. *et al.* A global model of natural volatile organic compound emissions. *Journal of Geophysical Research: Atmospheres* **100**, 8873-8892 (1995).
13 Ehn, M. *et al.* A large source of low-volatility secondary organic aerosol. *Nature* **506**, 476-479 (2014).
14 Bianchi, F. *et al.* New particle formation in the free troposphere: A question of chemistry and timing. *Science* **352**, 1109-1112 (2016).
15 Bianchi, F. *et al.* Highly oxygenated organic molecules (HOM) from gas-phase autoxidation involving peroxy radicals: A key contributor to atmospheric aerosol. *Chemical reviews* **119**, 3472-3509 (2019).
16 Tröstl, J. *et al.* The role of low-volatility organic compounds in initial particle growth in the atmosphere. *Nature* **533**, 527-531 (2016).
17 Vereecken, L., Müller, J. F. & Peeters, J. Low-volatility poly-oxygenates in the OH-initiated atmospheric oxidation of α-pinene: impact of non-traditional peroxyl radical chemistry. *Physical Chemistry Chemical Physics* **9**, 5241-5248 (2007).
18 Vereecken, L. & Peeters, J. Nontraditional (per) oxy ring-closure paths in the atmospheric oxidation of isoprene and monoterpenes. *The Journal of Physical Chemistry A* **108**, 5197-5204 (2004).
19 Yao, L. *et al.* Atmospheric new particle formation from sulfuric acid and amines in a Chinese megacity. *Science* **361**, 278-281 (2018).
20 Nie, W. *et al.* Secondary organic aerosol formed by condensing anthropogenic vapours over China's megacities. *Nature Geoscience* **15**, 255-261 (2022).
21 Guo, S. *et al.* Remarkable nucleation and growth of ultrafine particles from vehicular exhaust. *Proceedings of the National Academy of Sciences* **117**, 3427-3432 (2020).
22 Hoffmann, T. *et al.* Formation of organic aerosols from the oxidation of biogenic hydrocarbons. *Journal of Atmospheric Chemistry* **26**, 189-222 (1997).

23 Atkinson, R. & Arey, J. Gas-phase tropospheric chemistry of biogenic volatile organic compounds: a review. *Atmospheric Environment* **37**, 197-219 (2003).
24 Sindelarova, K. *et al.* Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmospheric Chemistry and Physics* **14**, 9317-9341 (2014).
25 Lee, A. *et al.* Gas‐phase products and secondary aerosol yields from the ozonolysis of ten different terpenes. *Journal of Geophysical Research: Atmospheres* **111** (2006).
26 Iyer, S. *et al.* Molecular mechanism for rapid autoxidation in α-pinene ozonolysis. *Nature communications* **12**, 878 (2021).
27 Zhang, X. *et al.* Formation and evolution of molecular products in α-pinene secondary organic aerosol. *Proceedings of the National Academy of Sciences* **112**, 14168-14173 (2015).
28 Quéléver, L. L. *et al.* Effect of temperature on the formation of highly oxygenated organic molecules (HOMs) from alpha-pinene ozonolysis. *Atmospheric Chemistry and Physics* **19**, 7609-7625 (2019).
29 Rissanen, M. P. *et al.* Effects of chemical complexity on the autoxidation mechanisms of endocyclic alkene ozonolysis products: From methylcyclohexenes toward understanding α-pinene. *The Journal of Physical Chemistry A* **119**, 4633-4650 (2015).
30 Kurten, T. *et al.* Computational study of hydrogen shifts and ring-opening mechanisms in α-pinene ozonolysis products. *The Journal of Physical Chemistry A* **119**, 11366-11375 (2015).
31 Kristensen, K. *et al.* High-molecular weight dimer esters are major products in aerosols from α-pinene ozonolysis and the boreal forest. *Environmental Science & Technology Letters* **3**, 280-285 (2016).
32 Meder, M. *et al.* Selective deuteration as a tool for resolving autoxidation mechanisms in α-pinene ozonolysis. *Atmospheric Chemistry and Physics* **23**, 4373-4390 (2023).
33 Saunders, S. M., Jenkin, M. E., Derwent, R. & Pilling, M. Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds. *Atmospheric Chemistry and Physics* **3**, 161-180 (2003).
34 Aumont, B., Szopa, S. & Madronich, S. Modelling the evolution of organic carbon during its gas-phase tropospheric oxidation: development of an explicit model based on a self generating approach. *Atmospheric Chemistry and Physics* **5**, 2497-2517 (2005).
35 Camredon, M., Aumont, B., Lee-Taylor, J. & Madronich, S. The SOA/VOC/NO x system: an explicit model of secondary organic aerosol formation. *Atmospheric Chemistry and Physics* **7**, 5599-5610 (2007).
36 Rissanen, M. P. *et al.* The formation of highly oxidized multifunctional products in the ozonolysis of cyclohexene. *Journal of the American Chemical Society* **136**, 15596-15606 (2014).
37 Sandström, H., Rissanen, M., Rousu, J. & Rinke, P. Data‐Driven Compound Identification in Atmospheric Mass Spectrometry. *Advanced Science* **11**, 2306235 (2024).
38 Wang, L. *et al.* Discovering chemistry with an ab initio nanoreactor. *Nature chemistry* **6**, 1044-1048 (2014).
39 Grimme, S. Exploration of chemical compound, conformer, and reaction space with meta-dynamics simulations based on tight-binding quantum chemical calculations. *Journal of chemical theory and computation* **15**, 2847-2862 (2019).

40 Raucci, U., Rizzi, V. & Parrinello, M. Discover, sample, and refine: Exploring chemistry with enhanced sampling techniques. *The Journal of Physical Chemistry Letters* **13**, 1424-1430 (2022).
41 Yang, M. *et al.* Combined molecular dynamics and coordinate driving method for automatic reaction pathway search of reactions in solution. *Journal of Chemical Theory and Computation* **14**, 5787-5796 (2018).
42 Unsleber, J. P. & Reiher, M. The exploration of chemical reaction networks. *Annual review of physical chemistry* **71**, 121-142 (2020).
43 Xu, R., Meisner, J., Chang, A. M., Thompson, K. C. & Martínez, T. J. First principles reaction discovery: from the Schrodinger equation to experimental prediction for methane pyrolysis. *Chemical Science* **14**, 7447-7464 (2023).
44 Chang, A. M., Meisner, J., Xu, R. & Martínez, T. J. Efficient Acceleration of Reaction Discovery in the Ab Initio Nanoreactor: Phenyl Radical Oxidation Chemistry. *The Journal of Physical Chemistry A* **127**, 9580-9589 (2023).
45 Pietrucci, F. & Andreoni, W. Graph theory meets ab initio molecular dynamics: atomic structures and transformations at the nanoscale. *Physical review letters* **107**, 085504 (2011).
46 Invernizzi, M. & Parrinello, M. Exploration vs convergence speed in adaptive-bias enhanced sampling. *Journal of Chemical Theory and Computation* **18**, 3988-3996 (2022).
47 Das, S., Raucci, U., Neves, R. P., Ramos, M. J. & Parrinello, M. How and when does an enzyme react? Unraveling α-Amylase catalytic activity with enhanced sampling techniques. *ACS Catalysis* **13**, 8092-8098 (2023).
48 Raucci, U., Sanchez, D. M., Martínez, T. J. & Parrinello, M. Enhanced sampling aided design of molecular photoswitches. *Journal of the American Chemical Society* **144**, 19265-19271 (2022).
49 Klippenstein, S. J. & Elliott, S. N. OH Roaming during the Ozonolysis of α-Pinene: A New Route to Highly Oxygenated Molecules? *The Journal of Physical Chemistry A* **127**, 10647-10662 (2023).
50 Xu, L., Møller, K. H., Crounse, J. D., Kjaergaard, H. G. & Wennberg, P. O. New insights into the radical chemistry and product distribution in the OH-initiated oxidation of benzene. *Environmental Science & Technology* **54**, 13467-13477 (2020).
51 Møller, K. H., Kurten, T., Bates, K. H., Thornton, J. A. & Kjaergaard, H. G. Thermalized epoxide formation in the atmosphere. *The Journal of Physical Chemistry A* **123**, 10620-10630 (2019).
52 Iyer, S., Rissanen, M. P. & Kurtén, T. Reaction between peroxy and alkoxy radicals can form stable adducts. *The journal of physical chemistry letters* **10**, 2051-2057 (2019).
53 Taatjes, C. A. *et al.* Hydroxyacetone production from C3 Criegee intermediates. *The Journal of Physical Chemistry A* **121**, 16-23 (2017).
54 Liu, T. *et al.* OH roaming and beyond in the unimolecular decay of the methyl-ethyl-substituted Criegee intermediate: observations and predictions. *Journal of the American Chemical Society* **145**, 19405-19420 (2023).
55 Liu, T. & Lester, M. I. Roaming in the Unimolecular Decay of syn-Methyl-Substituted Criegee Intermediates. *The Journal of Physical Chemistry A* **127**, 10817-10827 (2023).
56 Zádor, J. *et al.* Automated reaction kinetics of gas-phase organic species over multiwell potential energy surfaces. *The Journal of Physical Chemistry A* **127**, 565-588 (2023).

57 Teng, A. P., Crounse, J. D. & Wennberg, P. O. Isoprene peroxy radical dynamics. *Journal of the American Chemical Society* **139**, 5367-5377 (2017).
58 Wu, D. & Bayes, K. D. Rate constants for the reactions of isobutyl, neopentyl, cyclopentyl, and cyclohexyl radicals with molecular oxygen. *International journal of chemical kinetics* **18**, 547-554 (1986).
59 Atkinson, R. Gas-phase tropospheric chemistry of volatile organic compounds: 1. Alkanes and alkenes. *Journal of Physical and Chemical Reference Data* **26**, 215-290 (1997).
60 Zang, H. *et al.* Direct probing of acylperoxy radicals during ozonolysis of α-pinene: constraints on radical chemistry and production of highly oxygenated organic molecules. *Atmospheric Chemistry and Physics* **23**, 12691-12705 (2023).
61 Vereecken, L. & Nozière, B. H migration in peroxy radicals under atmospheric conditions. *Atmospheric chemistry and physics* **20**, 7429-7458 (2020).
62 Otkjær, R. V., Jakobsen, H. H., Tram, C. M. & Kjaergaard, H. G. Calculated hydrogen shift rate constants in substituted alkyl peroxy radicals. *The Journal of Physical Chemistry A* **122**, 8665-8673 (2018).
63 Hansen, A. S. *et al.* Watching a hydroperoxyalkyl radical (• QOOH) dissociate. *Science* **373**, 679-682 (2021).
64 O'Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *Journal of cheminformatics* **3**, 1-14 (2011).
65 Bonomi, M. *et al.* Promoting transparency and reproducibility in enhanced molecular simulations. *Nature methods* **16**, 670-673 (2019).
66 Kühne, T. D. *et al.* CP2K: An electronic structure and molecular dynamics software package-Quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics* **152** (2020).
67 Stewart, J. J. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular modeling* **13**, 1173-1213 (2007).
68 Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of chemical physics* **126** (2007).
69 Halgren, T. A. MMFF VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular ‐ interaction energies and geometries. *Journal of Computational Chemistry* **20**, 730-748 (1999).
70 Beck, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys* **98**, 5648-5646 (1993).
71 Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical review B* **37**, 785 (1988).
72 Hehre, W. J., Ditchfield, R. & Pople, J. A. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics* **56**, 2257-2261 (1972).
73 Frisch, M. J., Pople, J. A. & Binkley, J. S. Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *The Journal of chemical physics* **80**, 3265-3269 (1984).
74 Chai, J. D. & Head Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics* **10**, 6615-6620 (2008).

75 Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of chemical physics* **90**, 1007-1023 (1989).
76 Kendall, R. A., Dunning, T. H. & Harrison, R. J. Electron affinities of the first‐row atoms revisited. Systematic basis sets and wave functions. *The Journal of chemical physics* **96**, 6796-6806 (1992).
77 Watts, J. D., Gauss, J. & Bartlett, R. J. Coupled-cluster methods with noniterative triple excitations for restricted open-shell Hartree-Fock and other general single determinant reference functions. Energies and analytical gradients. *The Journal of chemical physics* **98**, 8718-8733 (1993).
78 Adler, T. B., Knizia, G. & Werner, H. J. A simple and efficient CCSD (T)-F12 approximation. *The Journal of chemical physics* **127** (2007).
79 Knizia, G., Adler, T. B. & Werner, H.-J. Simplified CCSD (T)-F12 methods: Theory and benchmarks. *The Journal of chemical physics* **130** (2009).
80 Werner, H. J., Knizia, G. & Manby, F. R. Explicitly correlated coupled cluster methods with pair-specific geminals. *Molecular Physics* **109**, 407-417 (2011).
81 Peterson, K. A., Adler, T. B. & Werner, H. J. Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B–Ne, and Al–Ar. *The Journal of chemical physics* **128** (2008).
82 Werner, H. J. *et al.* The Molpro quantum chemistry package. *The Journal of chemical physics* **152** (2020).
83 Glowacki, D. R., Liang, C. H., Morley, C., Pilling, M. J. & Robertson, S. H. MESMER: an open-source master equation solver for multi-energy well reactions. *The Journal of Physical Chemistry A* **116**, 9545-9560 (2012).
84 Shannon, R. J., Robertson, S. H., Blitz, M. A. & Seakins, P. W. Bimolecular reactions of activated species: An analysis of problematic HC (O) C (O) chemistry. *Chemical Physics Letters* **661**, 58-64 (2016).
85 Rissanen, M. P., Mikkilä, J., Iyer, S. & Hakala, J. Multi-scheme chemical ionization inlet (MION) for fast switching of reagent ion chemistry in atmospheric pressure chemical ionization mass spectrometry (CIMS) applications. *Atmospheric Measurement Techniques* **12**, 6635-6646 (2019).