

Molecular Dynamics-Guided Reaction Discovery Reveals Endoperoxide-to-Alkoxy Radical Isomerization as Key Branching Point in α -pinene Ozonolysis

Corresponding Author: Professor Theo Kurtén

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Significance and impact

The Yang et al. manuscript claims two principal contributions to our understanding of highly oxygenated organic molecule (HOM) formation that in turn form secondary organic aerosols (SOAs) and have important impact on climate and human health.

The first contribution is that they present a pathway, which first forms an endoperoxide, commonly found in these systems, and then transforms it into an intermediate with an alkoxy radical with an epoxy functional group. The pathway is found using MD simulations and is presented as “unexpected” and “newly discovered”. The second contribution is that they create an intricate kinetic model that includes chemical activation across several bimolecular steps, and show that the new pathway is likely relevant for HOMs. The paper and the claims are exclusively relying on theory and computations. Regarding the experiments, the authors sum it up as “Many of these routes likely contribute to the mass spectrometry peaks for C7 and C10 HOMs observed in our flow reactor experiments”, which is not a proof of the theory being right, rather, it is a weaker observation that the theory is not in conflict with the observations.

In our review related to significance and impact, we were interested to investigate the two key aspects of the manuscript, (a) the novelty of the paths, and (b) the soundness of the chemical activation theory.

Point (a) is relevant, because this is essentially the title, and thus the selling point of the paper. The authors mention two prior works describing the key step unveiled in this paper. Xu et al. looked at the endoperoxide rearrangement reaction in conjunction with the OH-initiated oxidation of benzene. The pathway was deemed unimportant in the atmosphere relative to O₂ addition, with a rate coefficient of 0.02/s. Xu et al. do not seem to provide the barrier height, but presumably it is relatively high because for the reaction to proceed the resonance stabilization needs to be broken. Importantly, one of the main conclusions of the Xu et al. paper is that this rearrangement reaction is not important, this is even highlighted in their TOC figure. Zador et al. looked at a non-resonance stabilized version of this reaction and got a significantly higher, 4200/s rate coefficient corresponding to a 13.1 kcal/mol barrier – this barrier is lower than the one described for α -pinene in this paper (17.9 kcal/mol), and falls into the category of lower barriers mentioned later in the manuscript. They pointed out the likely important role this channel can play in other systems, and also found seven decomposition channels for the product, which also agree with the findings of this paper: beta-scission to a carbonyl is the most prominent pathway, but several H-transfer reactions can play a role as well. Nevertheless, that paper was not focused on this reaction and did not follow up on the relevance of such reactions in a broader sense.

However, in light of these two papers, we wonder whether the authors can claim that this is a new pathway. Reading the title gives the impression that the MD-based reaction discovery tool found a new reaction with substantial impact. In contrast, it appears to us that the MD-based tool has not uncovered any fundamentally new (low-energy) pathways, and the potential strength of the paper is that it shows that the endoperoxide pathway that has been shown to exist before by others may have an important impact on HOM chemistry. This is related to the accuracy of the kinetic model and the thus topic (b) mentioned above. In our opinion a title such as “Endoperoxide to alkoxy isomerization plays a significant role in HOM chemistry” would

do better justice to the content, relevance, and novelty of the manuscript.

Regarding (b), the chemical model, there are positive and negative aspects of the work. The positive is that the authors were careful at investigating the various uncertainties in the applied model by looking at best and worst case scenarios, or lower and upper limits. Nevertheless, there are a lot of assumptions and approximations that go into this part, like model transition states, assumed rate coefficient ranges, etc. We do not think that any of the methods used are flawed, the authors are certainly experts in the field, but the conclusions are more speculative than definitive.

Technical comments

1. What is the computational cost of the MD-based exploration in CPU (or GPU) hours?
2. "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." How general/black box is the method? It is criticized (see next point) that others "rely heavily on chemical intuition", but a tailored CV in some ways can be similarly problematic (needs intuition) with important consequences as seen from the example.
3. "However, thus far, only experimental observation-based hypothesis-driven computational approaches that rely heavily on chemical intuition have been used to identify autoxidation pathways." It is a misrepresentation of the current trend. We think that the authors overlook the large body of recent literature in gas-phase reactions, including autoxidation, where theoretical techniques are moving away from being heavily relying on intuition. Current techniques are more discovery based, often with a reverse information flow to experiment, i.e., theory guides experiments.
4. Please define what a "model transition state" is. Is it an assumed barrier?
5. What is (on p11) pseudo-unimolecular? Do you mean pseudo-first-order?
6. Figure S6 shows concentration profiles up to about 1×10^{15} seconds, this is about 32 million years. Why? This is clearly not the relevant timescale.
7. The SI says: "For cases where O₂ addition were treated reversibly, the influence of the excess energy is minimal as large amounts of yields were caused by the reversible reaction of O₂." Given that the reversible assumption is the realistic one, we are somewhat baffled by this statement as it is buried here. Does it mean that the observed chemical outcomes, the core of the paper, are not due to excess energy? In contrast, the main paper says: "Lowering the rearrangement reaction barrier to approximately 14 kcal/mol is necessary for the back reaction to exert influence under realistic RO₂ lifetimes of 1 s to 100 s." This is for the same system, α -pinene.
8. The manuscript says: "Its ozonolysis reaction (α -pinene + O₃) 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 reported³²." The first sentence says that α -pinene + O₃ makes SOA. The second that the link is not conclusive. Which one is true? Please refine the statement.
9. "However, without an added dimension of separation³⁷, mass spectrometric techniques cannot distinguish different molecules with the same mass." What dimensions do the authors mean? There are broad ranges of MS techniques.
10. "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 RO₂ are far too slow to compete with bimolecular termination even in the cleanest conditions." Can you be clearer? In other parts of the paper, it is written: "It is noteworthy that, in contrast to previous results³⁰, the excess energy may also accelerate the H-shift reaction of certain first-generation RO₂ species derived from α -pinene, enabling the nanosecond-scale formation of some H-shift products (Supplementary Materials section 4)." Essentially, it is written in a somewhat confusing way what is rapid and what is not.
11. "In addition, by exploiting the analogy between this reaction and a previously reported epoxidation reaction⁵¹" What does exploiting mean here? Do the authors simply mean: "by analogy" or "using analogy"?
12. "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 + O₃ + nO₂ system)" It is not clear to us under what circumstances a pathway would not be real. The authors do not seem to mention IRCs for instance. Or multireference character. How likely it is for the applied method to produce non-physical pathways?
13. "Notably, the reaction discovery engine also independently identified the OH roaming pathway, which was recently proposed by Klippenstein and Elliott based on systematic theoretical calculations⁴⁹" Can you mark these in Fig. 2?
14. In the SI: "addition rate of EPO pat" Pat? – typo.
15. Please put page numbers in the SI.

Summary

- What are the noteworthy results?
 - o Showing that a certain endoperoxide rearrangement reaction may be important for HOM chemistry in the atmosphere.
- Will the work be of significance to the field and related fields?
 - o Yes, it is an interesting hypothesis and others might want to test it.
- How does it compare to the established literature? If the work is not original, please provide relevant references.
 - o We found no contradiction to previous works.
 - o However, many elements of the finding are not new. Notably, the invoked key pathway.
- Does the work support the conclusions and claims, or is additional evidence needed?
 - o By and large yes, but see technical comments. It is important to note, however, that there is no clear experimental evidence for the mechanism. Probably further work is needed.
- Are there any flaws in the data analysis, interpretation and conclusions? Do these prohibit publication or require revision?

o Generally not, but see technical comments.

- Is the methodology sound? Does the work meet the expected standards in your field?

o Overall yes, see technical comments.

- Is there enough detail provided in the methods for the work to be reproduced?

o Probably not. There are a lot of assumptions and specialized tools used. The geometries etc. are also not provided.

Reviewer #2

(Remarks to the Author)

Yang et al. provide a novel illustration of the utility of automated reaction discovery methods through an exploration of the pathways for multistage oxidation of alpha-pinene, the dominant biogenic volatile organic compound. This application highlights a novel reaction pathway involving the formation of alkoxy epoxide compounds as an interesting route to a wide variety of multiple oxygenated products, including new possible HOM precursors to secondary organic aerosols. The work is well described, state-of-the-art, and of interest to a general audience. I thoroughly enjoyed reading it. Thus, I highly recommend its publication, essentially as is. I have only a few very minor recommendations of places where additional discussion could be worthwhile.

The automated discovery calculations employ the PM6 level of theory. Many levels of theory would have challenges exploring roaming radical channels without some sort of forced inclusion of symmetry-breaking wavefunctions (to obtain unrestricted singlet states). Can the authors provide any discussion as to why such issues are not a problem at the PM6 level?

The authors properly emphasize the ability of their methodology to discover the roaming radical pathways of Klippenstein and Elliott, as illustrated in Fig. 2. However, there should also be roaming radical pathways from the VHP-b, which are not shown in Fig. 2. Does this mean that the automated discovery methodology did not find that pathway? If so, that should also be emphasized.

Another channel that one might suspect could exist would involve the formation of a diradical from the POZ as has been explored, for example, for $C_2H_4 + O_3$. It doesn't appear that the automated discovery approach discovered such a channel. Do the authors have any comments on why that might be?

The authors provide nice discussion of the dependence of the branching ratios on the somewhat uncertain rates for the $RB + O_2 = RB-O_2$ rate constant. It seems that it would be appropriate to also say a few words about the effect of corresponding variations/uncertainty in other $R + O_2$ rates, such as that for $VinOxy-a + O_2$ and $EPO + O_2$.

It would be helpful if Table S1 included relative energies in addition to the absolute energies. Additional tables listing the full sets of energies for other calculations reported in the manuscript would also be helpful. Similarly, some sort of tabulation of rate constants reported in the manuscript could be valuable to other researchers. In addition, inclusion of the MESMER input files as some sort of data attachment would greatly enhance the reproducibility of this effort.

One curiosity question I have is with respect to the presence of two plateaus in the $VinOxy-a$ concentration within Fig. S4 (and the related peak in the RB concentration). I presume that this behavior arises from some sort of back isomerization of RB . But I would have thought the RB and $VinOxy-a$ would simply slowly equilibrate without any sort of overproduction of RB followed by redissociation. Is this perhaps due to a change in the statistical branching between RB and $VinOxy-a$ with decreasing excess energy? If so, that would be interesting from a chemical physics perspective.

In Figure S6 it would be helpful to add the A, B, C, D labels to the plots themselves.

It might also be helpful to provide more detail on how the authors treated the presence of multiple conformers in their rate analysis, and also what kind of uncertainties might arise from whatever treatment is employed.

I found only two typos - in Section S5, it looks like Bill Gates managed to change "bimolecular" to "biomolecular". Also, after Table S2, "insensitivity" should be "insensitive".

Reviewer #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career

Researchers who co-review manuscripts.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I am satisfied with the answers.

Reviewer #2

(Remarks to the Author)

The authors have provided effective responses to all of my comments. I consider it appropriate for publication in nature communications.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Responses to reviewer #1 comments:

General comments: *The Yang et al. manuscript claims two principal contributions to our understanding of highly oxygenated organic molecule (HOM) formation that in turn form secondary organic aerosols (SOAs) and have important impact on climate and human health.*

The first contribution is that they present a pathway, which first forms an endoperoxide, commonly found in these systems, and then transforms it into an intermediate with an alkoxy radical with an epoxy functional group. The pathway is found using MD simulations and is presented as “unexpected” and “newly discovered”. The second contribution is that they create an intricate kinetic model that includes chemical activation across several bimolecular steps, and show that the new pathway is likely relevant for HOMs. The paper and the claims are exclusively relying on theory and computations. Regarding the experiments, the authors sum it up as “Many of these routes likely contribute to the mass spectrometry peaks for C7 and C10 HOMs observed in our flow reactor experiments”, which is not a proof of the theory being right, rather, it is a weaker observation that the theory is not in conflict with the observations.

In our review related to significance and impact, we were interested to investigate the two key aspects of the manuscript, (a) the novelty of the paths, and (b) the soundness of the chemical activation theory.

Point (a) is relevant, because this is essentially the title, and thus the selling point of the paper. The authors mention two prior works describing the key step unveiled in this paper. Xu et al. looked at the endoperoxide rearrangement reaction in conjunction with the OH-initiated oxidation of benzene. The pathway was deemed unimportant in the atmosphere relative to O₂ addition, with a rate coefficient of 0.02/s. Xu et al. do not seem to provide the barrier height, but presumably it is relatively high because for the reaction to proceed the resonance stabilization needs to be broken. Importantly, one of the main conclusions of the Xu et al. paper is that this rearrangement reaction is not important, this is even highlighted in their TOC figure. Zador et al. looked at a non-resonance stabilized version of this reaction and got a significantly higher, 4200/s rate coefficient corresponding to a 13.1 kcal/mol barrier – this barrier is lower than the one described for alpha-pinene in this paper (17.9 kcal/mol), and falls into the category of lower barriers mentioned later in the manuscript. They pointed out the likely important role this channel can play in other systems, and also found seven decomposition channels for the product, which also agree with the findings of this paper: beta-scission to a carbonyl is the most prominent pathway, but several H-transfer reactions can play a role as well. Nevertheless, that paper was not focused on this reaction and did not follow up on the relevance of such reactions in a broader sense.

However, in light of these two papers, we wonder whether the authors can claim that this is a new pathway. Reading the title gives the impression that the MD-based reaction discovery tool found a new reaction with substantial impact. In contrast, it appears to us that the MD-based tool has not uncovered any fundamentally new (low-energy) pathways, and the potential strength of the paper is that it shows that the endoperoxide pathway that has been shown to exist before by others may have an important impact on HOM chemistry. This is related to the accuracy of the kinetic model and the thus topic (b) mentioned above. In our opinion a title such as “Endoperoxyde to alkoxy isomerization plays a significant role in HOM chemistry” would do better justice to the content, relevance, and novelty of the manuscript.

Regarding (b), the chemical model, there are positive and negative aspects of the work. The

positive is that the authors were careful at investigating the various uncertainties in the applied model by looking at best and worst case scenarios, or lower and upper limits.

Nevertheless, there are a lot of assumptions and approximations that go into this part, like model transition states, assumed rate coefficient ranges, etc. We do not think that any of the methods used are flawed, the authors are certainly experts in the field, but the conclusions are more speculative than definitive.

General response: we thank the reviewers for their very thorough work, and their friendly constructive criticism, which has helped us substantially improve the manuscript. We agree that the endoperoxide rearrangement pathway has indeed been suggested previously, and have accordingly toned down our claims of novelty concerning this aspect of our work (as well as revised the title as suggested). However, we would argue – and the reviewers do not seem to disagree with us here – that the importance of the pathway is still major news, both for the specific case of alpha-pinene (where accurate treatment of the effect of excess energy plays an important role), and for atmospheric chemistry more generally (as illustrated by the low-barrier examples we provide beyond the alpha-pinene system). In terms of “proof of the theory being right”, we of course agree, as it is philosophically impossible to prove a theory or hypothesis conclusively. Instead, we suggest that the experimental evidence should be seen as an at least moderately stringent falsification test – which our hypothesis passes. While not conclusive, this is not insignificant either. Or in other words, while the hypothesis proposed here may not be the only possible explanation for the observed (previously unexplained) mass spectrometric (MS) peaks, it is arguably the best explanation presented so far. This is important, as MS remains the dominant experimental tool for studying atmospheric oxidation – but cannot, on its own, resolve actual molecular structures. Thus, complementary methods (including computational modelling such as that presented here) remain crucial tools for connecting MS peaks to specific structures.

We also suggest that, in addition to the principal contributions correctly identified by the reviewers, our study also makes a number of valuable and novel secondary contributions, including hitherto undiscovered roaming channels, as well as demonstrating that the reaction discovery method correctly reproduces all reaction pathways so far proposed in the literature (including some very recent and exotic ones). Especially the results on roaming have been expanded upon based on the comments of the second reviewer, who seems to agree with us that these results are also interesting to the community.

Concerning the statement that various assumptions and approximations make our conclusions somewhat speculative, we partially agree – but also point out that we carry out a number of additional test calculations in response to the specific technical comments by all the reviewers. (These include, among other things, yield estimation for the endoperoxide rearrangement pathway computed without model transition states, assessment of multi-conformer effects, as well as a brief discussion on multi-reference effects, as detailed further below). All of these help to increase the reliability of our results, and thus make our conclusions less speculative. Additionally, to ensure reproducibility of our results we have uploaded all the input files and optimized molecular structure files to Zenodo [DOI: 10.5281/zenodo.14006848].

Technical comments:

Point No. 1: What is the computational cost of the MD-based exploration in CPU (or GPU) hours?

Response No. 1: thank you for highlighting this point. The computational cost of the MD-based discovery simulations is comparable to that of regular MD simulations, and it is primarily determined by the level of theory employed. Specifically, the computational cost associated with evaluating the graph-based collective variable and its derivatives is minimal compared to the cost of evaluating the energy and forces required for the MD simulation. From this perspective, this methodology is highly generalizable and can be implemented using both GPU and CPU codes for running molecular dynamics. In our particular case, we used the CPU version of CP2K, and with 4 cores on a desktop workstation, we were able to perform each step in an average of 0.02 seconds. As a result, we were able to complete 500 ps of simulation in approximately 6 hours on average.

We have clarified this point in the Methods section.

Changes made in the manuscript (pages 18 & 19 lines 459-465):

“The computational cost of these MD-based discovery simulations is comparable to that of regular MD simulations, and it is primarily determined by the electronic level of theory. From this perspective, this methodology is highly generalizable and can be implemented using both GPU and CPU codes for running molecular dynamics. In our particular case, we used the CPU version of CP2K, and with 4 cores on a desktop workstation, we were able to perform each step in an average of 0.02 seconds. As a result, we were able to complete 500 ps of simulation in approximately 6 hours on average.”

Point No. 2: “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.” How general/black box is the method? It is criticized (see next point) that others “rely heavily on chemical intuition”, but a tailored CV in some ways can be similarly problematic (needs intuition) with important consequences as seen from the example.

Response No. 2: we agree that methods based on collective variables (CVs) often rely heavily on chemical intuition, such as when distances or combinations thereof are used as CVs. To address this issue, we previously proposed a more general approach to define a CV by representing a molecule as a graph. This method allows flexibility in choosing whether to include all atoms in the system (J. Phys. Chem. Lett. 2022, 13, 1424) or only the reactive functional groups (ACS Catal. 2023, 13, 8092) in the graph's definition.

In this paper, we approached the MD discovery process in a more general manner, treating it as a black-box tool by including **all atoms in the system** within the graph-based CV definition. However, this generality can present limitations, particularly for flexible molecules where the reaction is restricted to a specific functional group. In such cases, the graph-based CV explores a vast configurational space mainly dominated by conformational sampling rather than reaction pathways. Therefore, to observe the -OH breaking pathway of the vinyl hydroperoxide species

we redefined the graph to focus exclusively on the relevant functional group (-COOH), which proved to be a more efficient strategy in this case.

We acknowledge that this choice somewhat reduces the 'black box' nature of the method. However, we would argue that the reliance on chemical intuition is minimal since the regions excluded from the graph CV are primarily inert portions of the molecule, such as alkyl groups. We believe that focusing the graph CV on the reactive groups would not significantly compromise the generality of this discovery approach, particularly for atmospheric radical reactions where the 'reactive part' of the molecule is easily identifiable—typically near the radical center. Moreover, we would like to highlight that, among the numerous reactions discovered using our approach, only in this specific case we did tailor the graph CV to a 'reactive group'. In all other cases, every atom in the system was included in the CV.

Additionally, even though only the -COOH group was included in the graph CV, we still observed several products stemming from the roaming OH radical pathway, where the OH radical abstracts hydrogen atoms. Interestingly, these pathways were discovered in a "black-box" manner since the hydrogen atoms involved were not explicitly included in the CV definition. This underscores the point that while some general choices, based on established chemical rules, can be made to accelerate the discovery process, the resulting outcomes may deviate from those predicted by chemical intuition. Thus, the tool demonstrates its versatility and strength in revealing unexpected reaction pathways.

Changes made in the manuscript (page 18 lines 440-442):

“Notably, even with this focused CV, we still observed unexpected products, such as those from the roaming OH radical pathway, demonstrating the tool's capacity to discover reaction pathways beyond those predicted by chemical intuition.”

Point No. 3: *“However, thus far, only experimental observation-based hypothesis-driven computational approaches that rely heavily on chemical intuition have been used to identify autoxidation pathways.” It is a misrepresentation of the current trend. We think that the authors overlook the large body of recent literature in gas-phase reactions, including autoxidation, where theoretical techniques are moving away from being heavily relying on intuition. Current techniques are more discovery based, often with a reverse information flow to experiment, i.e., theory guides experiments.*

Response No. 3: we mostly agree, though with a few caveats. First, we would suggest splitting up the issue into two related but still separate trends: that of “intuition-driven research” vs “discovery” on one hand, and that of “theory guiding experiment” vs “experiment guiding theory” on the other. For the latter, we fully agree with the reviewer that this has been a trend already for a decade or more. Indeed, we are ourselves co-authors on a number of papers doing precisely this. However, in almost all of these studies, the theory guiding the experiments (for example, in terms of which reaction mechanisms to study) has still been largely based on chemical expertise and intuition. For the former case (intuition vs discovery), we also agree that a recent trend is emerging – but we emphasize that especially in the context of complex atmospheric reactions, fully automated discovery is very recent indeed (e.g. the Zador paper is only from last year). So while the present manuscript is not the first to attempt this (in the context of our field), it is one of the first papers. We have clarified this distinction also in the manuscript.

Changes made in the manuscript (page 3 lines 89-93):

“In the recent decades, computations have become efficient and accurate enough that theoretical studies are sometimes even able to guide experiments, rather than the other way around. However, except for a few very recent studies,^{38,39} both experimental and theoretical studies in the field of atmospheric reaction mechanisms have so far relied heavily on chemical expertise and intuition to identify, for example, autoxidation pathways.”

Point No. 4: Please define what a “model transition state” is. Is it an assumed barrier?

Response No. 4: it is correct. the model transition state has an assumed barrier that was not explicitly calculated. Optimizing transition states for O₂ addition reactions to alkyl radicals using single reference methods is difficult. However, a TS is nevertheless needed for the SimpleRRKM method in MESMER. While the assumed barrier was arbitrary, it was constrained by the rate of the O₂ addition, which was kept within the ambient relevant range of 10⁷ s⁻¹ to 10⁸ s⁻¹ as described in the manuscript.

We would like to clarify that the assumptions made regarding the model transition state do not directly affect the AOE yield. Instead, the yield is ultimately governed by the conversion rate of RB + O₂ to RB-RO₂. This rate controls the amount of excess energy retained in RB-RO₂—the faster the production rate, the less excess energy is dissipated. In turn, this retained excess energy influences whether subsequent isomerization steps are activated. This argument is further illustrated in the following example, where we estimate the AOE yield over the first 7 nanoseconds without resorting to the model transition state, assuming a moderate RB-RO₂ production rate of 7.1 × 10⁷ s⁻¹ and a moderate heat dissipation rate ($\Delta E_{\text{down}} = 175 \text{ cm}^{-1}$, Figure S2):

- In the first 2 nanoseconds, over 70% of RB was produced from the reaction of α -pinene with O₃, as determined by the MESMER simulation (Figure S4).
- In the following 5 nanoseconds (i.e., between 2 and 7 nanoseconds), the 70% of RB produced in the first 2 nanoseconds will result in at least 50% RB-RO₂, based on the calculation of a pseudo-first-order rate process: $70\% \times \{1 - \exp[-(5 \times 10^{-9} \text{ s}) \times (7.1 \times 10^7 \text{ s}^{-1})]\} = 50\%$.
- According to the ED-HD model shown in Figure S2, RB is expected to retain about 60 kcal/mol of excess energy at 7 nanoseconds. This implies that the 50% RB-RO₂ produced within this timeframe should retain at least (60 + 32.9) kcal/mol of excess energy (the 32.9 kcal/mol is from the RB + O₂ step).
- We performed additional MESMER simulations on a reduced PES starting from RB-RO₂ to determine the AOE yield, factoring in various levels of excess energy on RB-RO₂. This approach bypasses the use of the model transition state on the full PES. The results (Figure S13 and Table S5) indicate that with (60 + 32.9) kcal/mol of excess energy on RB-RO₂, the branching ratio between AOE and EPO-RO₂ is 59% vs. 41%.
- Therefore, based on points c and d, the estimated AOE yield over the first 7 nanoseconds is 50% × 59% = 29.5%. This value is quite close to the 26% yield (with an O₂ addition rate of 7.1 × 10⁷ s⁻¹) calculated in Figure 3 from the MESMER simulation on the full PES. The slightly higher yield in this estimation arises because the excess energy lost due to OH in the VHP decomposition step is not accounted for.

This estimation example utilizes a moderate RB-RO₂ production rate of 7.1 × 10⁷ s⁻¹ and a moderate heat dissipation rate ($\Delta E_{\text{down}} = 175 \text{ cm}^{-1}$, Figure S2). However, we have also explored

the lower and upper limits of both the RB-RO₂ production rate and the heat dissipation rate, as shown in Figure 3 and Figure S3 (using the full PES). We believe these simulations would underscore the significance of the newly identified pathway.

These have been clarified in the revised manuscript.

Changes made in the manuscript (page 10 lines 255-259):

“The SimpleRRKM method used to treat the O₂ addition step requires a transition state, but these are difficult to optimize. Therefore, we used a model transition state with the correct number of rovibrational modes but with an energy that was varied to provide a pseudo-first-order O₂ addition rate within the atmospherically relevant^{51,57,58} range of 10⁷ s⁻¹ to 10⁸ s⁻¹. ”

Changes made in the Supplementary Materials (pages 22, 23, & 24):

“In the main text Figure 3, we performed MESMER simulation on a full PES with a model transition state for the RB + O₂ to RB-RO₂ step. The model transition state has an assumed barrier that was not explicitly calculated. Optimizing transition states for O₂ addition reactions to alkyl radicals is difficult. However, a TS is nevertheless needed for the SimpleRRKM method in MESMER. While the assumed barrier was arbitrary, it was constrained by the rate of the O₂ addition, which was kept within the ambient relevant range of 10⁷ s⁻¹ to 10⁸ s⁻¹ as described in the manuscript.

We would like to clarify that the assumptions made regarding the model transition state do not directly affect the AOE yield. Instead, the yield is ultimately governed by the conversion rate of RB + O₂ to RB-RO₂. This rate controls the amount of excess energy retained in RB-RO₂—the faster the production rate, the less excess energy is dissipated. In turn, this retained excess energy influences whether subsequent isomerization steps are activated. This argument is further illustrated in the following example, where we estimate the AOE yield over the first 7 nanoseconds without resorting to the model transition state, assuming a moderate RB-RO₂ production rate of $7.1 \times 10^7 \text{ s}^{-1}$ and a moderate heat dissipation rate ($\Delta E_{\text{down}} = 175 \text{ cm}^{-1}$, Figure S2):

- a. In the first 2 nanoseconds, over 70% of RB is produced from the reaction of α -pinene with O₃, as determined by the MESMER simulation (Figure S4).*
- b. In the following 5 nanoseconds (i.e., between 2 and 7 nanoseconds), the 70% of RB produced in the first 2 nanoseconds will result in at least 50% RB-RO₂, based a pseudo-first-order rate process: $70\% \times \{1 - \exp[-(5 \times 10^{-9} \text{ s}) \times (7.1 \times 10^7 \text{ s}^{-1})]\} = 50\%$.*
- c. According to the ED-HD model shown in Figure S2, RB is expected to retain about 60 kcal/mol of excess energy at 7 nanoseconds. This implies that the 50% RB-RO₂ produced within this timeframe should retain at least (60 + 32.9) kcal/mol of excess energy (the 32.9 kcal/mol is from the RB + O₂ step).*
- d. We performed additional MESMER simulations on a reduced PES (Figure S13) starting from RB-RO₂ to determine the AOE yield, factoring in various levels of excess energy on RB-RO₂. This approach bypasses the use of the model transition state on the full PES. The results (Table S5) indicate that with (60 + 32.9) kcal/mol of excess energy on RB-RO₂, the branching ratio between AOE and EPO-RO₂ is 59% vs. 41%.*
- e. Therefore, based on points c and d, the estimated AOE yield over the first 7 nanoseconds is $50\% \times 59\% = 29.5\%$. This value is quite close to the 26% yield (with an O₂ addition rate of $7.1 \times 10^7 \text{ s}^{-1}$) calculated in Figure 3 from the MESMER simulation on the full PES. The slightly higher yield in this estimation arises because the excess energy lost due to OH in the VHP decomposition step is not accounted for.*

This estimation example utilizes a moderate RB-RO₂ production rate of $7.1 \times 10^7 \text{ s}^{-1}$ and a moderate heat dissipation rate ($\Delta E_{\text{down}} = 175 \text{ cm}^{-1}$, Figure S2). The lower and upper limits of both the RB-RO₂ production rate and the heat dissipation rate were explored and shown in Figure 3 and Figure S3 (using the full PES). These simulations underscore the significance of the newly identified pathway.

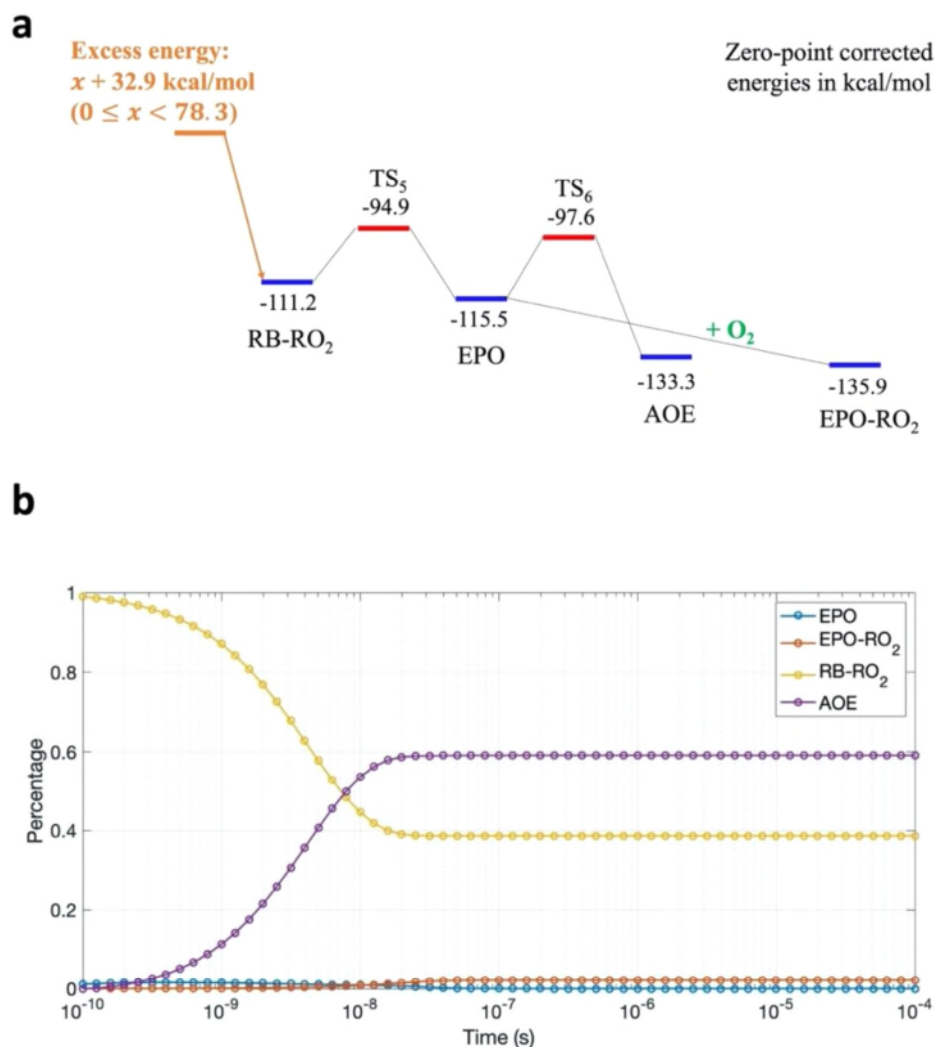


Figure S13. (A) Representation of the reduced potential energy surface employed in the kinetic master equation simulation to investigate the influence of excess energy on the AOE yield. (B) Temporal evolution of concentration profiles observed in the kinetic master equation simulation for the case of Total Excess Energy = 60 kcal/mol + 32.9 kcal/mol, where 60 kcal/mol is the assumed amount from the “ α -pinene to RB” step (whose maximum amount is 78.3 kcal/mol) and 32.9 kcal/mol comes from the “RB to RB-RO₂”. In this case, the amount of excess energy from the “ α -pinene to RB” step needs to be assumed because it is expected to be partially dissipated during the RB + O₂ step.

Table S5. Effect of the total excess energy possessed by RB-RO₂ on the yield of AOE

Total Excess Energy, kcal/mol (“ α -pinene to RB” + “RB to RB-RO ₂ ”)	AOE yield	EPO-RO ₂ yield
30 + 32.9	0.076	0.924
40 + 32.9	0.198	0.802

50 + 32.9	0.381	0.619
60 + 32.9	0.594	0.406

”

Point No. 5: What is (on p11) pseudo-unimolecular? Do you mean pseudo-first-order?

Response No. 5: we are indeed referring to pseudo-first-order. We have updated the terminology throughout the paper to ensure consistency. Thank you for bringing this to our attention.

Changes made in the manuscript:

- page 10 lines 257-261:

“Therefore, we used a model transition state with variable energies (while the O₂ concentration was fixed at $5 \times 10^{18} \text{ cm}^{-3}$) in simulations to control the effective pseudo-first-order O₂ addition rate. We found that, at an effective pseudo-first-order O₂ addition rate of $1.9 \times 10^7 \text{ s}^{-1}$ ”

- page 12 lines 292-294:

“The substantial yield observed within the plausible pseudo-first-order O₂ 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.”

- page 20 lines 502-504:

“employing a model transition state with variable energies (while O₂ concentration was fixed at $5 \times 10^{18} \text{ cm}^{-3}$) to modulate the effective pseudo-first-order addition rate within the range of about 10^7 s^{-1} to 10^8 s^{-1} ”

Changes made in the Supplementary Materials (pages 11, 16, & 17):

“we used a model transition state with variable energies (while the O₂ concentration was fixed at $5 \times 10^{18} \text{ cm}^{-3}$) to control the effective pseudo-first-order O₂ addition rate at $7.5 \times 10^7 \text{ s}^{-1}$.”

“we used a model transition state with variable energies (while the O₂ concentration was fixed at $5 \times 10^{18} \text{ cm}^{-3}$) to control the effective pseudo-first-order O₂ addition rate of EPO at either $0.8 \times 10^7 \text{ s}^{-1}$ or $1.2 \times 10^8 \text{ s}^{-1}$.”

Point No. 6: Figure S6 shows concentration profiles up to about 1×10^{15} seconds, this is about 32 million years. Why? This is clearly not the relevant timescale.

Response No. 6: we have adjusted the time scale in Figure S6 and Figure S7 to a more relevant range, now showing concentration profiles up to 1×10^6 seconds. We have also removed panel D of Figure S6 in the revised version due to the extended RO₂ lifetime.

Changes made in the Supplementary Materials (Figure S6 & S7):

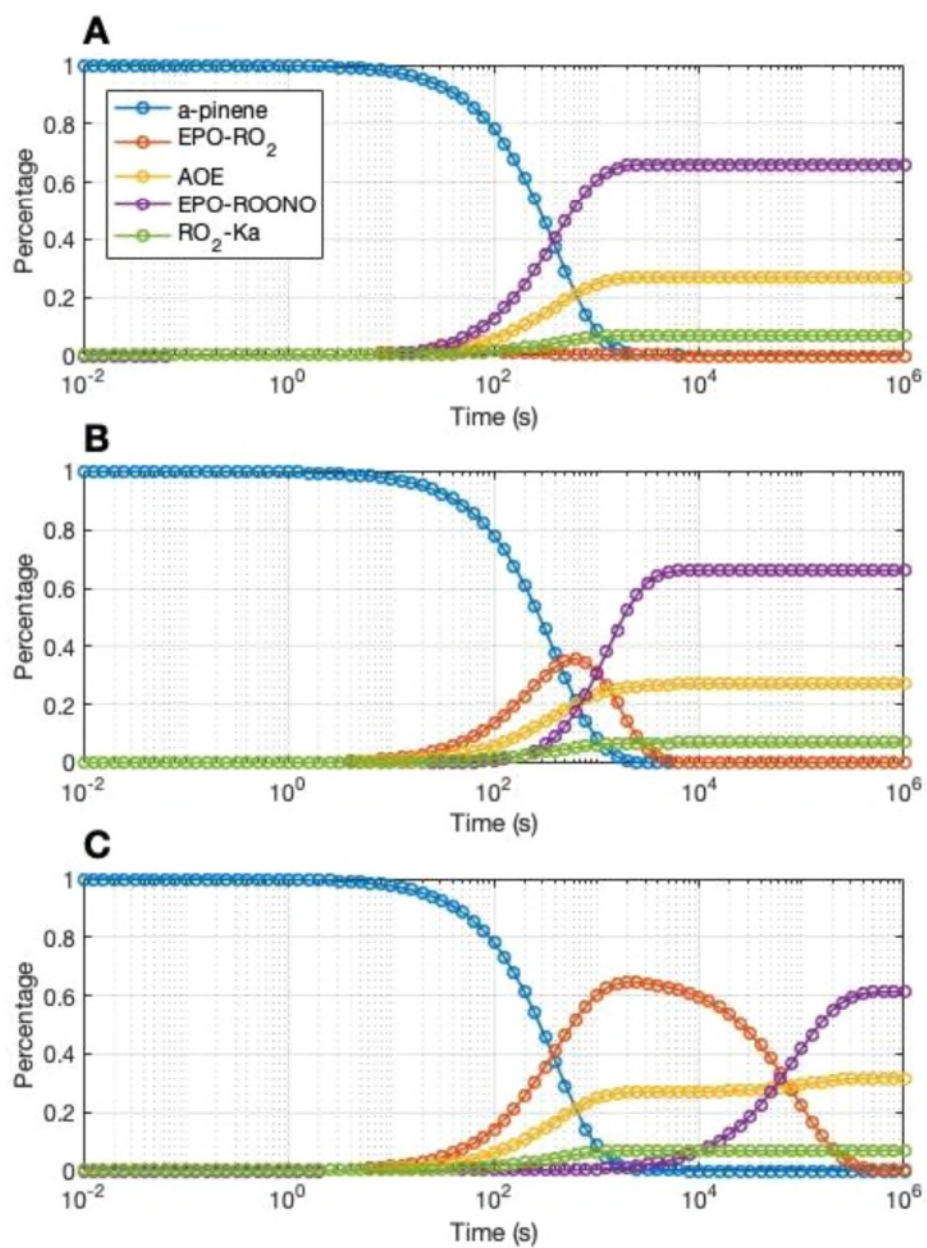


Figure S6. Temporal concentration profiles using different EPO-RO₂ lifetime. Simulations using the complete alpha-pinene ozonolysis PES in Figure 3 of the main text, but with different EPO-RO₂ lifetimes controlled by the bimolecular sink EPO-ROONO: (A) EPO-RO₂ lifetime = 10 s, (B) 10³ s, and (C) 10⁵ s.

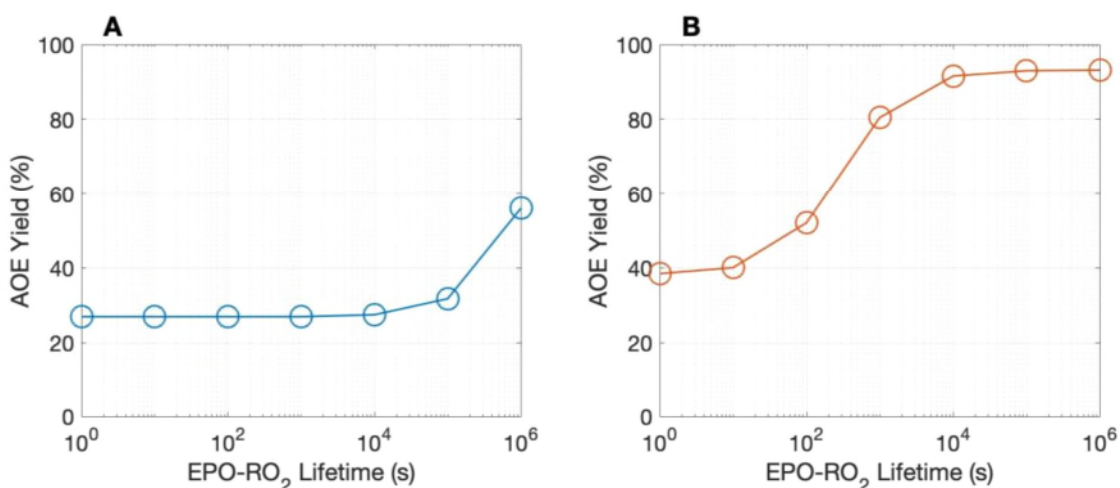


Figure S7. AOE yield as a function of EPO-RO₂ lifetime. (A) Simulations using the complete alpha-pinene ozonolysis PES in Figure 3 of the main text. (B) Simulations using the PES in Figure 3 of the main text, but with TS6 energy adjusted to be 14 kcal/mol.

Point No. 7: The SI says: “For cases where O₂ addition were treated reversibly, the influence of the excess energy is minimal as large amounts of yields were caused by the reversible reaction of O₂.” Given that the reversible assumption is the realistic one, we are somewhat baffled by this statement as it is buried here. Does it mean that the observed chemical outcomes, the core of the paper, are not due to excess energy? In contrast, the main paper says: “Lowering the rearrangement reaction barrier to approximately 14 kcal/mol is necessary for the back reaction to exert influence under realistic RO₂ lifetimes of 1 s to 100 s.” This is for the same system, alpha-pinene.

Response No. 7: thank you for your comment. We would like to clarify this point:

The chemical outcomes discussed in the main text, specifically the yield of the endoperoxy-alkyl radical rearrangement reaction, are entirely due to the excess energy. The differing conclusions between the main text and the SI arise from the analysis of different barrier heights for the endoperoxy-alkyl radical rearrangement (TS6) and varying amounts of excess energy.

In the main text, we modeled the full potential energy surface (PES) for alpha-pinene + O₃, where the TS6 barrier height is 17.9 kcal/mol, with an available excess energy of 115.5 kcal/mol (though a portion of the energy will dissipate before it can take effect). Because of the high excess energy value, the yield of the endoperoxy-alkyl radical rearrangement (AOE) is solely attributed to excess energy in the case of the main text. On the other hand, due to the high TS6 barrier, the EPO-RO₂ will primarily be consumed by bimolecular sink reactions before it can revert and cross TS6 to yield AOE. Lowering TS6 to around 14 kcal/mol is essential for the back reaction to impact AOE yield.

In contrast, the SI involves simulations on a reduced PES, where TS6 is varied between 6 and 14 kcal/mol, and the excess energy for EPO is set to either 0 or 20 kcal/mol. This section aims to provide context for the endoperoxy-alkyl radical rearrangement reactions that may occur in

other atmospheric VOCs, which likely have lower rearrangement barriers and lower excess energy from previous steps. We would like to clarify that although this set of simulations targets other atmospheric VOCs, the PESs were constructed by referencing several alpha-pinene intermediates for convenience. Our findings indicate that in scenarios with a maximum excess energy of 20 kcal/mol, the influence of reversible O₂ addition is significantly more pronounced than that of excess energy. The effects of excess energy is noticeable only under irreversible conditions.

We have revised the presentation to better differentiate the cases discussed in the main text and the SI.

Changes made in the Supplementary Materials (pages 16 & 17):

“In the main text, we modeled the full potential energy surface (PES) for alpha-pinene + O₃, where the TS6 barrier height is 17.9 kcal/mol, with an available excess energy of 115.5 kcal/mol (though a portion of the energy will dissipate before it can take effect). Because of the high excess energy value, the yield of the endoperoxy-alkyl radical rearrangement (AOE) is solely attributed to excess energy in the case of the main text. On the other hand, due to the high TS6 barrier, the EPO-RO₂ will primarily be consumed by bimolecular sink reactions before it can revert and cross TS6 to yield AOE. Lowering TS6 to around 14 kcal/mol is essential for the back reaction to impact AOE yield.

In this section, we performed additional simulations on a reduced PES, where TS6 barrier height is varied between 6 and 14 kcal/mol, and the excess energy for EPO is set to either 0 or 20 kcal/mol. This section aims to provide context for the endoperoxy-alkyl radical rearrangement reactions that may occur in other atmospheric compounds, which likely have lower rearrangement barriers and lower excess energy from previous steps. We found that the contributions of "excess energy" and the "back-reaction of O₂ addition" differ significantly in these cases compared to the full PES scenario discussed in the main text.

Although simulations performed in this section target other atmospheric compounds, the PESs were constructed by referencing several alpha-pinene intermediates for convenience.”

“The results are summarized in Figure S9. For these simulations on a reduced PES, where the TS6 barrier height varies between 6 kcal/mol to 14 kcal/mol and the excess energy for EPO is set to either 0 or 20 kcal/mol, we found that cases where O₂ addition was treated reversibly produced much higher yields than those treated irreversibly. Additionally, the influence of reversible O₂ addition was significantly more pronounced than that of excess energy. The effects of excess energy were only noticeable under irreversible conditions.

Overall, with no excess energy, the endoperoxy-alkyl radical rearrangement reaction is competitive to O₂ addition only if the reaction barrier is less than 8 kcal/mol – 10 kcal/mol, however, with 20 kcal/mol of excess energy, the barrier for the endoperoxy-alkyl radical rearrangement reaction to be competitive is elevated to 12 kcal/mol – 13 kcal/mol. This barrier height range is relevant to large portion of the similar reactions discovered for other atmospherically relevant systems (see section 7 of this document).”

Point No. 8: *The manuscript says: “Its ozonolysis reaction (α -pinene + O₃) 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 reported³².” The first sentence says that α -pinene + O₃ makes SOA. The second that the link is not conclusive. Which one is true? Please refine the statement.*

Response No. 8: both statements are correct but address different aspects of the α -pinene ozonolysis to SOA. The first sentence highlights that the α -pinene + O₃ system is indeed efficient for secondary organic aerosol (SOA) formation, a well-established fact. The second sentence, however, points out that while we know this reaction produces SOA, the exact pathways and mechanisms linking α -pinene ozonolysis to the formation of HOMs (the precursors of SOA) remain inconclusive.

To clarify, the formation of SOA from α -pinene and ozone can be viewed as a two-step process:

1. Autoxidation of α -pinene: This produces a range of oxygenated compounds, including HOMs with multiple oxygen atoms.
 2. HOM clustering and condensation: These HOMs either cluster with organic/inorganic species or condense on seed particles, contributing to SOA formation or increasing its mass.
- While chamber and field studies consistently identify HOMs formed from α -pinene + O₃ reactions using mass spectrometry (in both gas-phase and particulate forms), the challenge lies in resolving detailed molecular structures. As a result, the precise chemical mechanisms for how α -pinene ozonolysis forms specific HOMs remain unclear.

We have refined the relevant sentences to improve clarity.

Changes made in the manuscript (page 3 lines 68-72):

“Its ozonolysis reaction (α -pinene + O₃) 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}, an explicit chemical mechanism that links the α -pinene ozonolysis to the formation of various lowest-volatility HOMs, which are SOA precursors, has not yet been established³².”

Point No. 9: “However, without an added dimension of separation³⁷, mass spectrometric techniques cannot distinguish different molecules with the same mass.” What dimensions do the authors mean? There are broad ranges of MS techniques.

Response No. 9: this was a rather vague statement intended to indicate that it is relatively easy to infer the existence of several isomers with many MS methods, yet without an authentic standard and special analysis methods it is still close to impossible to definitively identify a complex multifunctional organic compound. As an example, some of the most routine ways to separate isomers are by their differing ion-mobilities and elution properties in chromatography columns and coupled with distinct fragmentation patterns (e.g., MS/MS and higher power collision induced dissociation analysis and electron impact ionization patterns) can yield molecular identification. An aspect we routinely take advantage is the unique tendency to form select ion-molecular adducts with several different reagent ions in chemical ionization analysis. We have now extended the text to clarify what was originally meant.

Changes made in the manuscript (page 3 lines 85-87):

“However, without an added dimension of separation provided by, for example, ion mobility or chromatographic analysis, mass spectrometric techniques cannot distinguish different molecules with the same mass.³⁷”

Point No. 10: *“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 RO₂ are far too slow to compete with bimolecular termination even in the cleanest conditions.” Can you be clearer? In other parts of the paper, it is written: “It is noteworthy that, in contrast to previous results³⁰, the excess energy may also accelerate the H-shift reaction of certain first-generation RO₂ species derived from α -pinene, enabling the nanosecond-scale formation of some H-shift products (Supplementary Materials section 4).” Essentially, it is written in a somewhat confusing way what is rapid and what is not.*

Response No. 10: there is indeed some ambiguity in our original presentation. Essentially, what matters is whether the unimolecular H-shift rates of RO₂ are faster than those of the corresponding bimolecular termination reactions in the atmosphere (e.g., RO₂ + HO₂, RO₂ + NO_x). Faster H-shift reaction rates allow autoxidation to continue, enabling repeated additions of O₂ to the carbon chain. Otherwise, the formation of HOMs is halted. As the pseudo-first-order rates for bimolecular termination reactions of RO₂ in the atmosphere exceed 10⁻² s⁻¹ even in the cleanest conditions (10⁻² s⁻¹ to 10² s⁻¹ in typical atmospheric conditions), "rapid" in this context refers to a rate higher than 10⁻² s⁻¹ to 10² s⁻¹.

The “textbook” reaction sequence (gray region in Figure 2 of the main text) outlines 4 first-generation RO₂ products. However, Kurtén et al.³⁰ calculated all possible H-shift reactions for the first-generation RO₂ and found that each of the four RO₂ can, at most, undergo one H-shift reaction (with a rate at around 10⁻² s⁻¹ to 10⁻¹ s⁻¹), while all other possible H-shifts are significantly slower (ranging from 10⁻²⁶ s⁻¹ to 10⁻³ s⁻¹). Consequently, under atmospheric conditions, such H-shifts can only lead to the addition of one more O₂ to the first-generation RO₂; i.e., autoxidation terminates at HOMs with 6 O atoms. In contrast, HOMs containing more than 10 O atoms are frequently observed in experiments.

The "rate between 10⁻² s⁻¹ and 10⁻¹ s⁻¹" mentioned in the previous paragraph refers to the rate under thermalized conditions, where excess energy is not considered. In this paper, we found that with excess energy, the rate of some H-shifts in first-generation RO₂ can be significantly accelerated, leading to nanosecond-scale formation of certain H-shift products (Supplementary Materials section 4). However, this excess energy likely affects only one H-shift cycle and dissipates rapidly, meaning it is unlikely to influence subsequent cycles. As a result, autoxidation still terminates at HOMs with 6 O atoms, though the yield of the current cycle may increase due to the excess energy.

We have revised relevant paragraphs to clarify this.

Changes made in the manuscript:

- page 6 lines 154-159:

“Three of the four CI isomers (as shown in Figure 2) have rapid 1,4-hydrogen shift channels, leading to the formation of three vinyl hydroperoxides (VHPs). These VHPs subsequently release an OH radical, producing vinyloxy radicals—resonance-stabilized species with both alkyl and alkoxy radical characteristics. The vinyloxy radicals then add O₂, forming four distinct types of peroxy radicals (RO₂), commonly referred to as the first-generation RO₂ in α -pinene ozonolysis.”

- pages 6 & 7 lines 162-176:

“The “textbook” reaction sequence, depicted in gray in Figure 2, cannot explain the experimentally observed HOMs with more than 6 O atoms in α -pinene ozonolysis. This is because the H-shift reactions of the first-generation RO₂ are generally too slow to compete with potential bimolecular termination reactions (with NO_x or HO₂) in the atmosphere. Even under the cleanest conditions, these bimolecular termination reactions typically exhibit pseudo-first-order reaction rates exceeding 10⁻² s⁻¹ (10⁻² s⁻¹ to 10² s⁻¹ in typical atmospheric conditions). Kurtén et al.³⁰ calculated all possible H-shift reactions for the first-generation RO₂ and found that each of the four RO₂ can, at most, undergo one H-shift reaction (with a rate at around 10⁻² s⁻¹ to 10⁻¹ s⁻¹), while all other H-shifts are significantly slower (ranging from 10⁻²⁶ s⁻¹ to 10⁻³ s⁻¹). Consequently, under atmospheric conditions, such H-shifts can only lead to the addition of one more O₂ to the first-generation RO₂. In 2021, Iyer et al.²⁶ partially addressed this issue by proposing an alternative reaction pathway, in which one of the vinyloxy radicals undergoes a chemically activated ring-opening, followed by O₂ addition to form a new RO₂ radical. This RO₂ is capable of rapid unimolecular isomerization, outcompeting bimolecular termination reactions and facilitating several more steps of autoxidation.”

- page 10 lines 266-271:

“It is noteworthy that, although previous results³⁰ indicate that the fastest H-shift reaction rates for first-generation RO₂ range from 10⁻² to 10⁻¹ s⁻¹ under thermalized conditions, these rates could be significantly enhanced by the excess energy discussed in this paper, potentially allowing for nanosecond-scale formation of certain H-shift products (see Supplementary Materials, section 4). However, the excess energy likely only influences one H-shift cycle and is unlikely to persist into subsequent cycles due to rapid dissipation.”

Point No. 11: *“In addition, by exploiting the analogy between this reaction and a previously reported epoxidation reaction⁵¹” What does exploiting mean here? Do the authors simply mean: “by analogy” or “using analogy”?*

Response No. 11: The term "exploiting" refers to utilizing or taking advantage of the analogy between the two reactions. We agree that "using analogy" or "by analogy" would convey the intended message more directly and have revised the sentence accordingly.

Changes made in the manuscript (page 5 lines 132-133):

“In addition, using the analogy between this reaction and a previously reported epoxidation reaction⁵¹”

Point No. 12: *“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 + O₃ + nO₂ system)” It is not clear to us under what circumstances a pathway would not be real. The authors do not seem to mention IRCs for instance. Or multireference character. How likely it is for the applied method to produce non-physical pathways?*

Response No. 12: we agree that the term “real” in this context is ambiguous, as is “non-physical”. After all, almost any combination of atomic motions physically possible, given sufficient energy. Of course, not all combinations of atomic motions will correspond to reaction pathways, i.e. minimum-energy pathways between two minima on the potential energy surface. Similarly, some combinations, while reaction pathways in this technical sense, will be so high in energy that they will never happen in reality, as something else (e.g. a simple decomposition

reaction) will inevitably happen first instead. And – as the reviewers presumably allude to when they mention multireference character– a combination of atomic motions that is a reaction pathway according to one level of theory may not be one according to another level of theory. Based on our experiences so far, it seems that the method employed here is surprisingly good (not perfect, but very good) at finding reaction channels that turn out to be actual reaction pathways also after re-optimization at a higher level of theory (including the IRCs mentioned by the reviewers) – albeit often with very different relative energetics, i.e., the relative energy ordering of the discovered pathways may be very different after the higher-level optimizations. Whether or not a reaction channel that turns out to have a very high barrier when reoptimized with higher-level methods should then be defined as “non-physical” or not is ultimately a matter of taste. As these distinctions are somewhat irrelevant to the present paper (where we are trying to discover and chart the channels that actually happen in atmospheric conditions), we have simply reformulated the sentence, making a less strong claim, and also adding a definition.

Additionally, we would like to briefly address the multi-reference effect. In our previous study (Nat. Commun., 2021, 12, 878; Supplementary Materials, first section), we optimized the reactant, transition state, and product of the vinoxy-a ring-breaking pathway using extended multi-configuration quasi-degenerate perturbation theory at second order (XMC-QDPT2). These optimizations were performed with eight electrons in six molecular orbitals (MOs) and seven electrons in six MOs, respectively. The calculations, carried out using Firefly software with the 6-311++G(d,p) basis set, yielded an XMC-QDPT2 barrier for the vinoxy-a ring-breaking reaction (Figure 3, TS4 in the main manuscript) that is 12.1 kcal/mol. This is a good verification of the ROHF-ROCCSD(T)-F12a/VDZ-F12//wB97X-D/aug-cc-pVTZ level of theory used in this study, where TS4 is 14.0 kcal/mol above the vinoxy-a intermediate.

Changes made in the manuscript (page 8 lines 199-202):

“While these reaction channels may represent actual reaction pathways (minimum-energy paths connecting potential energy surface minima) that could be accessible under suitable conditions, the vast majority of them are not competitive with bimolecular sinks in the atmosphere”

Point No. 13: *“Notably, the reaction discovery engine also independently identified the OH roaming pathway, which was recently proposed by Klippenstein and Elliott based on systematic theoretical calculations⁴⁹” Can you mark these in Fig. 2?*

Response No. 13: thank you for the suggestion. As also suggested by the other reviewer, we have revised Fig. 2 to better highlight the OH roaming pathways.

Changes made in the manuscript (Fig. 2):

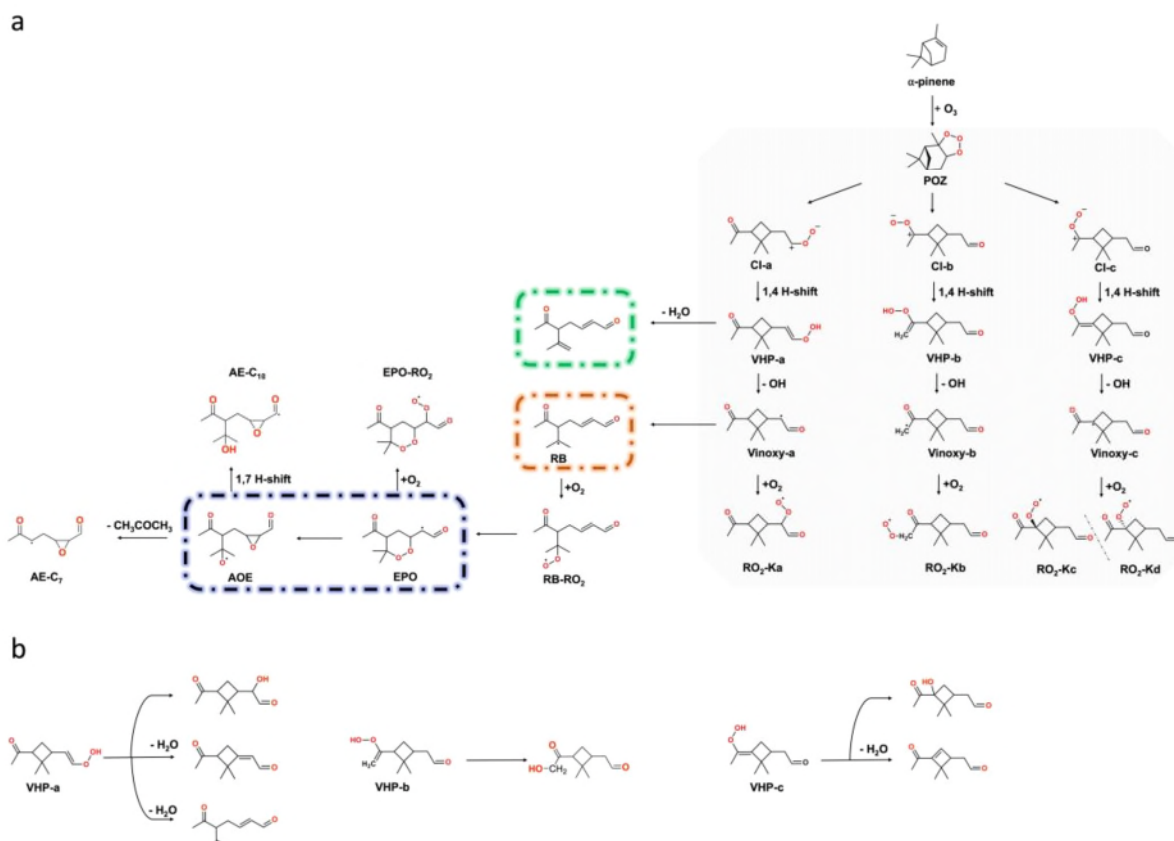


Figure 2. α -pinene ozonolysis mechanisms from the MD driven reaction discovery simulations. (a). The known and newly discovered pathways. The grey region represents the conventional perception of the α -pinene ozonolysis mechanisms. The ring-breaking mechanism proposed by Iyer et al.²⁶ is marked orange. 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. (b). Complete set of products from the OH roaming mechanism.

Point No. 14: In the SI: “addition rate of EPO pat” Pat? – typo.

Response No. 14: Thank you for pointing this out. We have corrected the typo. The sentence should read: "addition rate of EPO at either $0.8e7\text{ s}^{-1}$ or $1.2e8\text{ s}^{-1}$."

Changes made in the Supplementary Materials (page 17):

“addition rate of EPO at either $0.8e7\text{ s}^{-1}$ or $1.2e8\text{ s}^{-1}$.”

Point No. 15: Please put page numbers in the SI.

Response No. 15: Thank you for your suggestion. We have added page numbers to the Supplementary Materials.

Changes made in the Supplementary Materials: *page numbers have been added.*

Responses to reviewer #2 comments:

General comments: Yang et al. provide a novel illustration of the utility of automated reaction discovery methods through an exploration of the pathways for multistage oxidation of alpha-pinene, the dominant biogenic volatile organic compound. This application highlights a novel reaction pathway involving the formation of alkoxy epoxide compounds as an interesting route to a wide variety of multiple oxygenated products, including new possible HOM precursors to secondary organic aerosols. The work is well described, state-of-the-art, and of interest to a general audience. I thoroughly enjoyed reading it. Thus, I highly recommend its publication, essentially as is. I have only a few very minor recommendations of places where additional discussion could be worthwhile.

General response: We thank the reviewer for their very thorough evaluation and constructive suggestions, which have been invaluable in improving our manuscript. We are especially grateful for the reviewer's enthusiastic and highly favorable assessment of our work. Their recognition of its novelty, broad relevance, and potential impact further encourages us. Below, we provide a point-by-point response to each of the reviewer's comments.

Technical comments:

Point No. 1: *The automated discovery calculations employ the PM6 level of theory. Many levels of theory would have challenges exploring roaming radical channels without some sort of forced inclusion of symmetry-breaking wavefunctions (to obtain unrestricted singlet states). Can the authors provide any discussion as to why such issues are not a problem at the PM6 level?*

Response No. 1: thank you for pointing this out. To explore the roaming radical channel, we performed the dynamics with the system treated as an unrestricted singlet state, allowing us to break the symmetry of the wavefunction. This approach enabled the PM6 simulations to reveal the OH roaming pathway successfully.

Additionally, we have performed new MD discovery simulations at the unrestricted wB97xd/6-31G level, starting from VHP-a in Fig. 2. Using the same CV (including the COOH group in the graph), we have identified the same products as those obtained with the PM6 level of theory. Since these calculations rely on an *ab initio* Hamiltonian and are computationally more demanding, we employed a more aggressive bias deposition strategy, adding bias every 50 steps of the dynamics. We performed 11 simulations, varying the OPES explorer barrier value from 50 to 300 kJ/mol in increments of 25 kJ/mol, sampling a cumulative time of approximately 30 ps.

We have clarified this point in the Methods section.

Changes made in the manuscript (page 18 lines 442-446):

"To explore these channels, we treated the system as an unrestricted singlet state in the dynamics. Additionally, we have performed MD discovery simulations at the unrestricted wB97xd/6-31G level, starting from VHP-a in Fig. 2. Using the same CV (including the COOH group in the graph), we have identified the same products as those obtained with the PM6 level of theory."

Point No. 2: The authors properly emphasize the ability of their methodology to discover the roaming radical pathways of Klippenstein and Elliott, as illustrated in Fig. 2. However, there should also be roaming radical pathways from the VHP-b, which are not shown in Fig. 2. Does this mean that the automated discovery methodology did not find that pathway? If so, that should also be emphasized.

Response No. 2: thank you for the suggestion. We have also discovered the OH roaming pathway for VHP-b, and as also suggested by the other reviewer, we have revised Fig. 2 to better highlight the OH roaming pathways.

Changes made in the manuscript (Fig.2):

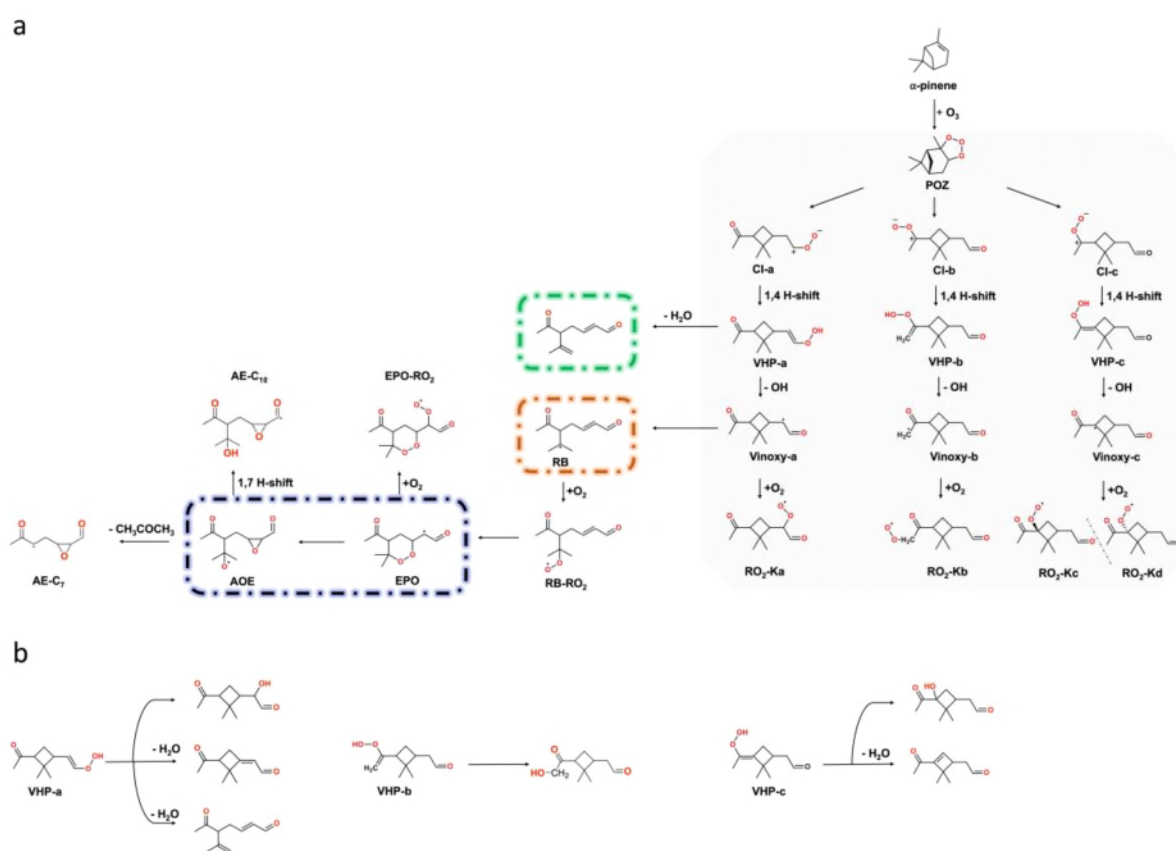


Figure 2. α -pinene ozonolysis mechanisms from the MD driven reaction discovery simulations. (a). The known and newly discovered pathways. The grey region represents the conventional perception of the α -pinene ozonolysis mechanisms. The ring-breaking mechanism proposed by Iyer et al.²⁶ is marked orange. 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. (b). Complete set of products from the OH roaming mechanism.

Point No. 3: Another channel that one might suspect could exist would involve the formation of a diradical from the POZ as has been explored, for example, for $C_2H_4 + O_3$. It doesn't appear that the automated discovery approach discovered such a channel. Do the authors have any comments on why that might be?

Response No. 3: We have not observed the formation of the diradical of POZ in the discovery simulations. A potential reason for this is that the step-wise diradical pathway for $C_2H_4 + O_3$ requires an internal C-C bond rotation to prevent the system from minimizing into the Criegee intermediate-forming transition state (as discussed in Nat. Comms. 2021, 12, 878, supporting information). In simpler systems, such a rotation is achievable. However, in the case of the α -pinene POZ system, the presence of a six-membered ring introduces significant steric hindrance, making C-C rotation highly restricted. This geometric constraint could be why we did not observe the diradical channel in the α -pinene system.

Point No. 4: The authors provide nice discussion of the dependence of the branching ratios on the somewhat uncertain rates for the $RB + O_2 = RB-O_2$ rate constant. It seems that it would be appropriate to also say a few words about the effect of corresponding variations/uncertainty in other $R + O_2$ rates, such as that for $Vinoxy-a + O_2$ and $EPO + O_2$.

Response No. 4: Thank you for the suggestion. In our kinetic master equation simulations performed on the complete α -pinene ozonolysis PES, we initially varied the pseudo-first-order rate constants for the three $R + O_2$ reactions ($RB + O_2$, $Vinoxy-a + O_2$, and $EPO + O_2$) simultaneously. This approach allowed us to assess their combined effects on the final yield of AOE.

To further address the individual impacts of the $Vinoxy-a + O_2$ and $EPO + O_2$ rates, we conducted additional simulations as described in the revised manuscript. In these simulations, we varied either the $Vinoxy-a + O_2$ or $EPO + O_2$ rate while holding the other two rates constant. The results indicate that variations in the $Vinoxy-a + O_2$ and $EPO + O_2$ rates have a negligible effect (<1%) on the AOE yield within the $1e7\text{ s}^{-1}$ to $1e8\text{ s}^{-1}$ range (see Supplementary Materials Section 12).

Changes made in the Supplementary Materials (page 25):

Table S6. Effect of the pseudo first-order O_2 addition rate for “ $Vinoxy-a + O_2$ ” on AOE yield. MESMER simulations were performed on the full potential energy surface shown in Figure 3 of the main text. The other two pseudo first-order O_2 addition rates for “ $EPO + O_2$ ” and “ $RB + O_2$ ” were fixed at $7.1 \times 10^7\text{ s}^{-1}$.

O_2 addition rate, s^{-1} ($Vinoxy-a + O_2$)	AOE yield	EPO- RO_2 yield	RO_2 -Ka yield
10^7	0.266	0.693	0.041
2×10^7	0.266	0.689	0.045
4×10^7	0.265	0.681	0.054
6×10^7	0.264	0.672	0.064
8×10^7	0.263	0.665	0.072
10^8	0.262	0.658	0.080

Table S7. Effect of the pseudo first-order O₂ addition rate for “EPO + O₂” on AOE yield. MESMER simulations were performed on the full potential energy surface shown in Figure 3 of the main text. The other two pseudo first-order O₂ addition rates for “Vinoxy-a + O₂” and “RB + O₂” were fixed at $7.1 \times 10^7 \text{ s}^{-1}$.

O ₂ addition rate, s ⁻¹ (EPO + O ₂)	AOE yield	EPO-RO ₂ yield	RO ₂ -Ka yield
10 ⁷	0.269	0.663	0.068
2×10 ⁷	0.268	0.664	0.068
4×10 ⁷	0.266	0.666	0.068
6×10 ⁷	0.265	0.667	0.068
8×10 ⁷	0.263	0.669	0.068
10 ⁸	0.261	0.671	0.068

Point No. 5: It would be helpful if Table S1 included relative energies in addition to the absolute energies. Additional tables listing the full sets of energies for other calculations reported in the manuscript would also be helpful. Similarly, some sort of tabulation of rate constants reported in the manuscript could be valuable to other researchers. In addition, inclusion of the MESMER input files as some sort of data attachment would greatly enhance the reproducibility of this effort.

Response No. 5: Thank you for your suggestions. We have now included relative energies for the molecules listed in Table S1 in the Supplementary Materials (Table S2). Additionally, the energies corresponding to the core pathways reported in the main text are summarized in Tables S8 and S9. All relevant rate coefficients for the studied cases are included in the MESMER output files (Table S10 in Point No. 8 as an example) and are now uploaded to Zenodo [DOI: 10.5281/zenodo.14006848]. To ensure the reproducibility of our work, we have also uploaded all MESMER input files, the molecular dynamics input files, and optimized molecular structure files to and Zenodo [DOI: 10.5281/zenodo.14006848].

Changes made in the Supplementary Materials (page 4):

Table S2. Relative energies in kcal/mol for molecules and transition states in Figure S1b. DFT = B3LYP/6-31+G(d); ZPE = Zero-point corrected energy; EE = Electronic energy

Molecules	DFT – EE	DFT – ZPE
1+2O ₂	0	0
2+2O ₂	6.54	5.59
3+O ₂	-26.26	-22.72
4+O ₂	-26.22	-23.38
5+O ₂	-35.46	-31.56
6	-53.32	-46.56
7	-50.86	-44.07
8+O ₂	-53.75	-51.48
9+O ₂	-18.60	-13.62
10+O ₂	-22.85	-17.46
11	-52.10	-43.14
12	-53.09	-43.94

13+O ₂	-73.00	-67.12
TS1+2O ₂	32.43	31.21
TS2+O ₂	-14.68	-12.32
TS3+O ₂	-15.09	-12.11
TS4+O ₂	-16.15	-14.56
TS5+O ₂	-8.70	-5.39
TS6+O ₂	-8.97	-4.84
TS7+O ₂	-5.85	-2.58
TS8+O ₂	-8.72	-4.85

Changes made in the Supplementary Materials (pages 26 & 27):

Table S8. Absolute energies in Hartrees for molecules shown in Figure 3 of the main text. ZPE = Zero-point corrected energy; EE = Electronic energy. Note that “ZP correction” is calculated at the wB97X-D/aug-cc-pVTZ level of theory.

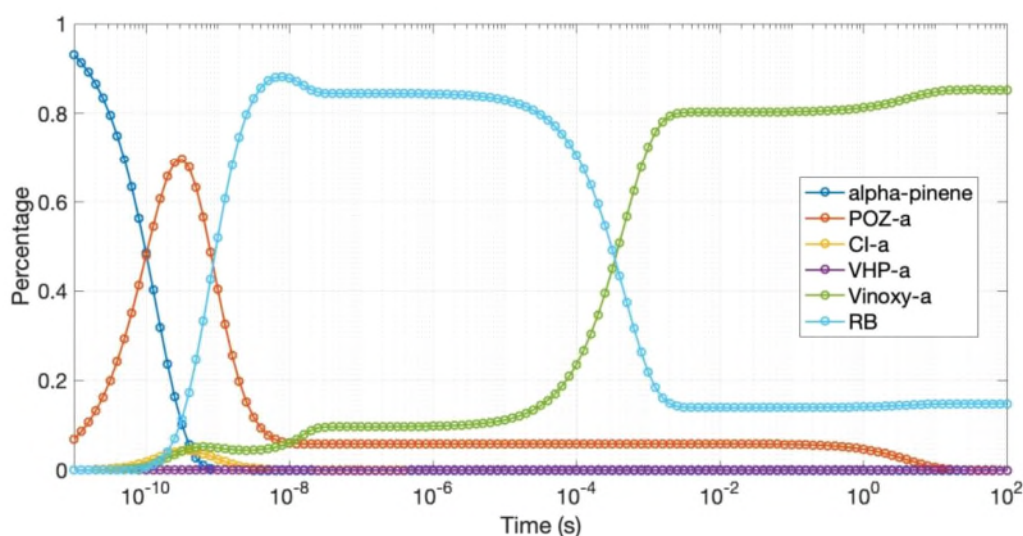
Molecules	ZP correction	CCSD(T) – E	CCSD(T) – ZPE
AP	0.237260	-390.0658103	-389.8285503
POZ	0.252126	-615.3704567	-615.1183307
CI-a	0.247038	-615.3930726	-615.1460346
VHP-a	0.246846	-615.4276444	-615.1807984
Vinoxy-a	0.229910	-539.7314706	-539.5015606
RO2-Ka	0.238798	-689.9454328	-689.7066348
RB	0.227132	-539.7242497	-539.4971177
RB-RO ₂	0.237990	-689.9590600	-689.7210700
EPO	0.239707	-689.9676416	-689.7279346
AOE	0.238038	-689.9943714	-689.7563334
EPO-RO ₂	0.249011	-840.1809889	-839.9319779
TS ₁	0.247577	-615.2637538	-615.0161768
TS ₂	0.248896	-615.3371465	-615.0882505
TS ₃	0.242706	-615.3631863	-615.1204803
TS ₄	0.227312	-539.7065330	-539.4792210
TS ₅	0.237363	-689.9325584	-689.6951954
TS ₆	0.237206	-689.9365995	-689.6993935
O ₂	0.003884	-150.1754180	-150.1715340
OH	0.008603	-75.66287400	-75.65427100
O ₃	0.007811	-225.2059050	-225.1980940

Table S9. Relative energies in kcal/mol for molecules shown in Figure 3 of the main text.
ZPE = Zero-point corrected energy; EE = Electronic energy

Molecules	CCSD(T) – EE	CCSD(T) – ZPE
<i>α-pinene+O₃+2O₂</i>	0	0
<i>POZ+2O₂</i>	-62.0	-57.5
<i>CI-a+2O₂</i>	-76.2	-74.9
<i>VHP-a+2O₂</i>	-97.9	-96.7
<i>Vinoxy-a+2O₂+OH</i>	-77.0	-81.1
<i>RO2-Ka+O₂+OH</i>	-101.1	-102.1
<i>RB+2O₂+OH</i>	-72.4	-78.3
<i>RB-RO2+O₂+OH</i>	-109.7	-111.2
<i>EPO+O₂+OH</i>	-115.1	-115.5
<i>AOE+O₂+OH</i>	-131.9	-133.3
<i>EPO-RO2+OH</i>	-138.9	-135.9
<i>TS1+2O₂</i>	5.0	6.6
<i>TS2+2O₂</i>	-41.1	-38.7
<i>TS3+2O₂</i>	-57.4	-58.9
<i>TS4+2O₂+OH</i>	-61.3	-67.0
<i>TS5+O₂+OH</i>	-93.1	-94.9
<i>TS6+O₂+OH</i>	-95.6	-97.6

Point No. 6: One curiosity question I have is with respect to the presence of two plateaus in the Vinoxy-a concentration within Fig. S4 (and the related peak in the RB concentration). I presume that this behavior arises from some sort of back isomerization of RB. But I would have thought the RB and Vinoxy-a would simply slowly equilibrate without any sort of overproduction of RB followed by redissociation. Is this perhaps due to a change in the statistical branching between RB and Vinoxy-a with decreasing excess energy? If so, that would be interesting from a chemical physics perspective.

Response No. 6: This is a very interesting point – we are not completely sure why this specific behaviour arises. However, the timescale for the shift (from the first Vinoxy-a plateau to the second) coincides with the timescale for full thermalization in the ED-HD model used in MESMER (as illustrated in Figure S2). Qualitatively, the direction of the change is the expected one (relatively less RB at lower excess energy). So the reviewer’s interpretation is probably the correct one: this is about the branching between the two species with decreasing excess energy. Having said that, the timescale for actual thermal equilibration will be considerably longer than that shown in Figure S4, as the thermal rate associated with a unimolecular reaction with a barrier of about 11 kcal/mol (the RB to TS energy difference) is on the order of 10^5 s^{-1} . Therefore, the plateau seen at the right-hand edges of Figure S4 are not yet the actual thermal equilibrium values, but the distribution that the system is “locked into” after all excess energy is lost. (Case-in-point: the thermal equilibrium concentration of the POZ will certainly be very close to zero, yet Figure S4 still shows some percent remaining of the POZ at 10^{-6} seconds). For your reference, we have attached a figure computed using the same setup as Figure S4, but with an extended time duration.



Point No. 7: *In Figure S6 it would be helpful to add the A, B, C, D labels to the plots themselves.*

Response No. 7: Thank you for the suggestion. We have updated Figure S6 to include the A, B, and C labels on the plots (We opted to remove panel D in the revised version due to the extended RO₂ lifetime).

Changes made in the Supplementary Materials (page 13): We have added labels to Figure S6.

Point No. 8: *It might also be helpful to provide more detail on how the authors treated the presence of multiple conformers in their rate analysis, and also what kind of uncertainties might arise from whatever treatment is employed.*

Response No. 8: Thank you for the suggestion. We performed full conformational sampling to obtain the lowest-energy conformers of the different stationary points along the PES, but multiple conformers were not considered in our RRKM-ME simulations. MESMER does not account for multiple conformers explicitly, but the use of the HinderedRotorQM1D method allows for the implicit consideration of conformers. To assess effects of multiple conformers, we used Gaussian 16 to run DFT calculations (wB97X-D/aug-cc-pVTZ) with Freq=HinderedRotor to identify the hindered rotors + the periodicities and barriers of all “non-sink” molecules along the PES in Figure 3. We successfully identified hindered rotors for all intermediates and transition states except TS₅. Hindered rotor calculations with Gaussian is imperfect and often fails to identify the internal rotational modes of complex transition states, such as TS₅. The following table compares the rate coefficients derived from MESMER for the rigid-rotor harmonic oscillator (RRHO) model versus the hindered rotor model. As shown, the rate coefficients are quite similar between the two cases (within a factor of ~2), except for the EPO → AOE step, where the hindered rotor model yields a rate coefficient approximately one order of magnitude higher than the RRHO model (the higher value favors the proposed endoperoxide rearrangement pathway). We also compared the fraction of alpha-pinene molecules that follow the Vinox-a → RB path (as opposed to the Vinox-a → RO₂-Ka path), which contributes to the subsequent endoperoxide rearrangement pathway. Using MESMER

simulations with pseudo-first-order O₂ addition rate coefficients of 7.1e7 s⁻¹ (RRHO) and 8.2e7 s⁻¹ (hindered rotor), we found similar results: 93.2% for the RRHO model and 96.8% for the hindered rotor model. These comparable values suggest that including or excluding multiple conformer effects will not alter our conclusions.

The related MESMER input files are included in the Zenodo [DOI: 10.5281/zenodo.14006848] data archive published along with this paper.

Table S10. MESMER derived rate coefficients at thermalized conditions: RRHO v.s. hindered rotor.

	RRHO (s ⁻¹)	Hindered rotor (s ⁻¹)
AP → POZ	2.043e-5	4.484e-5
POZ → CI-a	0.144	0.142
CI-a → VHP-a	299.67	173.37
VHP-a → Vinox-a	1470.06	1420.20
Vinox-a → RB	282.73	122.72
EPO → AOE	0.882	7.938

Changes made in the Manuscript (page 20 lines 519 - 521):

“Although our MESMER simulations did not account for multi-conformer effects, we have included a discussion of these effects in Section 13 of the Supplementary Materials.”

Changes made in the Supplementary Materials (page 28):

“13. Multi-conformer effects

We performed full conformational sampling to obtain the lowest-energy conformers of the different stationary points along the PES, but multiple conformers were not considered in our RRKM-ME simulations. MESMER does not account for multiple conformers explicitly, but the use of the HinderedRotorQM1D method allows for the implicit consideration of conformers. To assess effects of multiple conformers, we used Gaussian 16 to run DFT calculations (wB97X-D/aug-cc-pVTZ) with Freq=HinderedRotor to identify the hindered rotors + the periodicities and barriers of all “non-sink” molecules along the PES in Figure 3. We successfully identified hindered rotors for all intermediates and transition states except TS₅. Hindered rotor calculations with Gaussian is imperfect and often fails to identify the internal rotational modes of complex transition states, such as TS₅. The following table compares the rate coefficients derived from MESMER for the rigid-rotor harmonic oscillator (RRHO) model versus the hindered rotor model. As shown, the rate coefficients are quite similar between the two cases (within a factor of ~2), except for the EPO → AOE step, where the hindered rotor model yields a rate coefficient approximately one order of magnitude higher than the RRHO model (the higher value favors the proposed endoperoxide rearrangement pathway). We also compared the fraction of alpha-pinene molecules that follow the Vinox-a → RB path (as opposed to the Vinox-a → RO₂-Ka path), which contributes to the subsequent endoperoxide rearrangement pathway. Using MESMER simulations with pseudo-first-order O₂ addition rate coefficients of 7.1e7 s⁻¹ (RRHO) and 8.2e7 s⁻¹ (hindered rotor), we found similar results: 93.2% for the RRHO model and 96.8% for the hindered rotor model. These comparable values suggest that including or excluding multiple conformer effects will not alter our conclusions.”

Table S10. MESMER derived rate coefficients at thermalized conditions: RRHO v.s. hindered rotor.

	RRHO (s^{-1})	Hindered rotor (s^{-1})
$AP \rightarrow POZ$	2.043e-5	4.484e-5
$POZ \rightarrow CI-a$	0.144	0.142
$CI-a \rightarrow VHP-a$	299.67	173.37
$VHP-a \rightarrow Vinox-y-a$	1470.06	1420.20
$Vinox-y-a \rightarrow RB$	282.73	122.72
$EPO \rightarrow AOE$	0.882	7.938

Point No. 9: I found only two typos - in Section S5, it looks like Bill Gates managed to change “bimolecular” to “biomolecular”. Also, after Table S2, “insensitivity” should be “insensitive”.

Response No. 9: Thank you for your careful review and for bringing these typos to our attention. We have corrected the typos.

Changes made in the Supplementary Materials (pages 6 & 13):

“The Thermal Accommodation Coefficient (TAC) has been extensively studied through experiments and molecular dynamics simulations and was found to be insensitive to the temperature.”

“(2) We added a bimolecular “ $EPO-RO_2 + NO \rightarrow EPO-ROONO$ ” sink reaction to control the lifetime of $EPO-RO_2$ (varied between 1 s to 10^8 s)”