Derivation of Atmospheric Reaction Mechanisms for Volatile Organic Compounds by the SAPRC Mechanism Generation System (MechGen)

W.P.L. Carter, J. Jiang, J. J. Orlando, K. C. Barsanti
Manuscript ID: egusphere-2023-2343

We wish to thank the reviewers for taking the time to review this manuscript and their generally positive responses and their support for its publication. We especially wish to thank Dr. Luc Vereecken for taking the time to critically review all aspects of this manuscript and also the detailed proofreading. Although he also supports publication of this paper, Dr. Vereecken pointed out several areas where our estimates are not consistent with his understanding of the literature and results of work he has published. We agree with Dr. Vereecken that these are areas that need to be re-examined the next time that MechGen is updated. However, the many areas of chemistry that MechGen has to cover are subjects of active ongoing research where our knowledge is constantly evolving. Developing all of the SARs and estimates needed to derive complete atmospheric chemical mechanisms takes many years of work, and it is impossible to keep all aspects completely up to date at any given time. If that were required for publication then this could never be published in the foreseeable future, and users of MechGen would not be able to cite documentation to support the scientific basis of MechGen when they use it for their work. Instead this manuscript is intended to provide a snapshot of the system as currently exists, to allow it to be eventually used in published scientific studies. Dr. Vereecken recognizes this, and states that "this review is not the time to change how the mechanism generation is done". He primarily requests that we provide more citations of the literature and point out areas where updates are needed. We attempted to address these requests in the revised manuscript, and hope that he will agree that these are sufficient.

Given below are the reviewers' comments, followed by our responses and changes made to the manuscript where applicable. We will discuss the other reviewer comments first, before discussing the more detailed comments provided by Dr. Vereecken. Reviewer comments are indented and given in *italic font* to clearly distinguish them from our responses.

Comments by Tim Wallington

"The SAPRC mechanism has played a key role in atmospheric chemistry research and the development of policies to improve urban air quality since publication of the first version (SAPRC-90) in 1990. MechGen is a mechanism generation system which has been used to construct chemical schemes and estimate the rates of gas-phase chemical reactions and photolysis processes in the development of SAPRC mechanisms since 1999. This paper provides detailed documentation of the chemical basis of the version of MechGen which was used to construct the latest version of the SAPRC mechanism (SAPRC-22). The documentation is comprehensive and is presented in a clear and logical order. This is the first comprehensive peer-reviewed report describing MechGen and its publication will facilitate broader use of MechGen. It will be an important reference source both for the chemistry in MechGen and more broadly for the current state of atmospheric chemistry."

We thank Dr. Wallington for his positive comments, for which no reply is needed. Dr. Wallington is a well established expert in atmospheric chemistry and chemical kinetics, and has participated in or led a number of review panels regarding subjects covered by this paper. Therefore, his positive review is significant.

Comments by Reviewer 2

"As a key part of chemical transport modeling, gas phase mechanism is crucial in determining formation of secondary components, including both gases and particulate matter. However, there are no summarizing documents in recently years and many modelers are not familiar with the details. As the most detailed documentation of the SAPRC mechanism, I welcome its publication. Please find below comments for improving the manuscript.

1. The abstract just shows what are included in the manuscript, with key messages to give.
2. MechGen is the key program, however, this manuscript does not provide its basic idea, processes, and structure. The code of MechGen is also not provided.
3. The writing is too much like a report instead of scientific article, I would like to see more references to the statements."

We appreciate that the reviewer understands the importance of this work and supports its publication. Our responses to his specific comments are as follows:

1. 1. The "key message" is that this is the first peer-reviewed documentation of the scientific basis for SARs and estimates used by the MechGen system. We note that many of the SARs in the paper have not been previously published. The scientific significance of this paper is that this represents the first step in making MechGen more widely available for use by others and more consistent with FAIR data standards. Further, the manuscript provides a detailed roadmap for the community in seeing which areas of research are most needed to reduce uncertainties in atmospheric chemical mechanisms. Note that many other papers are published documenting SARs and estimates for mechanisms, such as the many papers by Jenkin and co-workers published in ACP. We have edited the abstract to make these points more clearly.

2. The purpose of this paper is to document the scientific basis of the chemical estimates used by MechGen, not the software itself. MechGen as a software system will be described in a subsequent publication that is still in preparation. The abstract was modified to point this out more clearly.

3. A number of additional references were added where needed. Most were in response to Dr. Vereecken's comments or suggestions, but additional citations were added as well when needed to support statements made in the text.

Comments by Luc Vereecken

"This manuscript describes the estimation methods for the atmospheric chemistry models generated by the current version of MechGen. As model generation must account for all reaction classes, the text covers a lot of topics, making this a very extensive document even with much of the discussion and tabulated material being presented in the supporting information. I support publication of the current work.

"In reviewing this work, I tried to separate between textual/presentation issues, the quality of the estimation methods within the chosen representation, and the choice of which chemistry to include and how to estimate it. I find the quality of the presentation of a high standard, with only small suggestions for improvement. The estimation methods as presented work well, and represent the state of the current knowledge within the chosen representation; I have no suggestions for improvement there. There are two topics (aromatics, and CI) where I feel the discussion/representation strays too far from how recent literature understands the underlying reaction mechanism. For the CI, this does not affect the final estimations, while for aromatics I think the representation is more problematic

(see below). Still, the manuscript describes the current state of the ever-evolving implementation of Mechgen very well, and this review is not the time to change how the mechanism generation is done. Many of the comments below are therefore mostly for additional literature data, or suggestions for future improvements rather than a critique of the current work, and need not be addressed other than perhaps an explicit mention in the text and/or SI to aid the reader (Note: I refer to own publications extensively below, not because they necessarily represent the best available but because it is easier for me to find suitable examples; citing other references may be more appropriate for this paper).

"A general comment is that the text relies often on "common knowledge" and is occasionally very light on references to the relevant literature. I recognize that this manuscript is not the place for an extensive literature review, but it in many places it uses "we assume" / "we estimate" without any literature provided, while one or two well-chosen references could provide a more solid foundation with provenance for the reader."

We appreciate Dr. Vereecken for taking the time to review this document in detail and provide helpful comments. Given below, we discuss his major substantive comments and criticisms first, followed by discussing his additional comments and edits. As indicated above, it is not possible to have a system as extensive as this be completely up to date at all times, and it is periodically necessary to temporarily freeze development so the system is available in a documented and citable form for use by the community. Once it is documented and published, our hope is to start working on an updated version that can be a subject of a future publication. Dr. Vereecken understands this and is not requesting that we modify MechGen at this time, but he is requesting that we point out the areas that may not be consistent with the latest literature, that will need to be the focus of these future updates. We agree with Dr. Vereecken that this needs to be discussed, and the manuscript was modified to more clearly point out these areas of potential conflict with the current literature and to add additional references where needed. However, this is not intended as a review paper, and no attempt was made to comprehensively review all the past literature on atmospheric chemical mechanisms. Instead the purpose is to document the estimates and assignments in the system, and to describe how they were derived.

Comments regarding Ozonolysis and Criegee Reactions

Section 2.1.4 VOC + O3 :
Designating all CI as excited when formed is not in agreement with the reaction mechanism: a fraction of stabilized CI are formed thermalized, especially for larger VOCs, with the remainder of the energy in the carbonyl fragment. There is no impact on the implementation, except that not separating between nascently thermalized CI, collisionally thermalized CI, and excited reactions of CI, requires a very different parametrization for the pressure-dependence. See also the related remarks below.

SI page 74: "Configuration is assumed not to be a factor for the excited intermediates because they are assumed to rapidly interconvert".
This is not in agreement with the ozonolysis mechanism as derived in many experimental and theoretical studies. The syn/anti isomerisation TS has barriers of over 25 kcal/mol even in the most favorable cases and 30+ kcal/mol being more common (Vereecken et al. 2017, 2022 and literature therein), such that syn/anti isomerisation has reaction rates several orders of magnitude below the various exit channels accessible to all CI with barriers significantly below that (at the very least 1,3 ring closure accessible to all CI with barrier = ~20 kcal/mol even in the worst case and <15 kcal/mol being more common). These other channels will be taken in preference, at any level of excitation attainable from VOC ozonolysis or carbene+O2 reactions. For all practical purposes the syn/anti ratio is set directly at the dissociation of the POZ. Additionally, in POZ dissociation there is a

3

fraction (small though it may be) of the CI that are formed stabilized, where "stabilized" means with an energy content comparable to or below the lowest exit channel barrier such that they will behave thermalized. The remaining energy from POZ dissociation is released towards the other fragment and the relative motion of the fragments. The yield of nascent stabilized CI depends on the size of the VOC, the relative size of the fragments with the concomitant probability distributions between the degrees of freedom, and the barriers for the available exit channels for the CI formed. The experimental yields of SCI as a function of pressure are difficult to reconcile with a generalized 100% excited CI formation especially for larger VOCs. Interesting here are the experiments by Drozd et al (2011), Newland et al (2020), and the predictive capabilities of the Newland et al SAR (2022) based on the energy distribution over the two fragments. Only for cyclic alkene ozonolysis and carbene+O2 reactions is it reasonable to assume 100% excited CI, as only a single product is formed.

In the end the Mechgen predictions for the product yields from VOC ozonolysis is not affected by how the CI formation and decomposition process is thought of, as the parametrization is designed to match the experimental observation irrespective of the underlying assumptions. Still, it might make sense to at least mention the possibility of a residual yield of stabilized CI even at low pressure, and state that syn/anti isomerisation has very high barriers compared to the other exit channels, making syn/anti isomerisation a negligible channel for any attainable excitation.

Line 1438: "In all cases they [CI] are expected to be formed in a vibrationally excited state, with the amount of excitation depending on the formation reaction, but sufficient to allow rapid interconversion between the syn and anti forms where applicable."

The treatment of the reactions of Criegee intermediates is one of the major areas that will need to be re-examined when MechGen is next updated. This will be a major effort, not only because of the remaining uncertainties that need to be dealt with, but also because some of the assumptions and parameters used is based on results of simulations of chamber experiments. As Vereecken noted, an update may not significantly affect predictions in many cases because they are constrained by experimental data. Therefore, we did not change the Criegee mechanisms, but pointed out the various issues in the discussion. The specific changes related to the substance of the Criegee discussion is as follows, with text in the revised manuscript in quotes, and changes made underlined (see also the manuscript and SI document where changes are highlighted).

Section 2.1.4, sentence added: "The assumption that the Criegee intermediates are formed entirely in the excited state is an approximation that is made to simplify the estimations of their subsequent reactions, as discussed in Section 7."

Section 7: Modifications to the middle of the first paragraph: "In all cases they [CI's] are expected to be formed in a vibrationally excited state, with the amount of excitation depending on the formation reaction, though the amount of excitation is expected to be variable because, except for reactions of O3 with cyclic compounds, the available energy is distributed among two fragments. In some cases the excitation may be sufficient to allow interconversions between syn and anti, though Vereecken et al. (2017, 2022, and references therein) calculated that the isomerization barriers are high, and interconversions of excited CIs may be slow in many cases. In any case, the interconversion …"

Section 7, new second paragraph: "For estimation purposes, we make the assumptions that CIs are initially formed in excited states and that the unimolecular reactions of the excited CIs do not depend on syn or anti configurations. Their configurations are determined only after they are stabilized. This is an oversimplification because the excited CIs may actually have different configurations and may not always interconvert rapidly compared to decompositions. However, available experimental and theoretical data are insufficient to derive the additional parameters required to make separate configuration-dependent estimates for excited as well as stabilized CIs. In many cases the estimates are constrained by

experimental or theoretical results, and this informs the parameters used in the estimates. Thus while the parameters may have been different if configuration were taken into account for excited CI reactions, the predictions would be similar. However, when applied to systems for which no data are available, the parameters and the predictions may be different, and are therefore more uncertain."

Section 7, middle of third paragraph: "Note that Vereecken et al. (2022) provided a more recent update on the reactions of oxygenated and unsaturated CIs, but these have not yet been taken into account in the current estimates."

> *Line 1515: "Excited intermediates with zero or one substituent can undergo "hot acid" rearrangements where an H bonded to the Criegee group is involved in a rearrangement forming a highly excited acid, which then decomposes in several ways"*
> *Decomposition of acids to anything other than CO, CO2 or ketene seems tenuous. as it is not supported by pyrolysis experiments of acids, not by theoretical literature which finds that all other pathways have too high a barrier to contribute significantly. It was a useful placeholder in early literature until better information was available, but other pathways have been found to explain some of the observed products. For example ethene ozonolysis (OH formation without a VHP channel) was often explained through hot acid decomposition, but seem more likely to be governed by ketohydroperoxide chemistry (see e.g. Pfeifle et al 2018 and later work on this channel). For some acid measurements it has been shown that they are probably secondary products (e.g. Orzechowska et al. 2005a/b).*

The following modifications of the paragraph discussing of "hot acid" rearrangements: "Excited intermediates with zero or one substituent can undergo "hot acid" rearrangements where an H bonded to the Criegee group is involved in a rearrangement forming a highly excited acid, which then decomposes in several ways, forming either radicals or stable compounds (e.g., reactions 1-5, 9-11, and 19-21 on Table 15). These reactions are assumed only to occur when the intermediate is excited. These estimates are based on data for simple alkenes and may not correctly predict products formed from more complex CIs with a single substituent. Other "hot acids" may not decompose in this way or may be mostly stabilized. These pathways will need to be re-examined when MechGen is updated."

> *Lines 1519-1526: ring closure of beta-carbonyl CI to a POZ*
> *The middle structure in the scheme requires a double bond between the carbon atoms, and is equivalent to the POZ formed in an alkyne+O3 reaction. Vereecken et al. 2017, and Cremer et al 2001 calculated a large barrier (= 30 kcal/mol) for this ring closure, and Vereecken et al. 2017 state explicitly that ringclosure is negligible against other CI reactions, making the O-rearrangement moot even for excited CI.*

> *[SI] page 82: "we assume that this type of O-shift reaction is faster than any competing possible reactions for excited carbonyl-substituted Criegees"*
> *This is in contradiction to the theoretical data, which finds unsurmountably high energy barriers for the ring closure to the POZ, compared to the other accessible reactions. (Cremer et al. 2001)*

The following was added to Section 7, where the reactions of excited CI's with α-carbonyl groups are discussed: "MechGen assumes that excited CIs with a-carbonyl groups are in rapid equilibrium with the primary ozonide and the isomer where the O is transferred to the other group, forming a different carbonyl-substituted intermediate if the groups bonded to them are different, i.e.:

 [revised to show double bond]
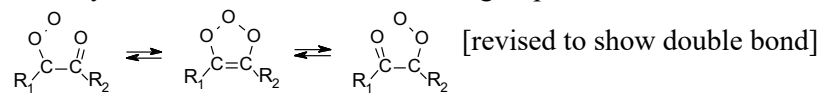
It further assumes the most reactive isomer determines the subsequent reactions, or that reaction of either form is equally likely if they undergo the same types of reactions. Although Vereecken et al. (2017) discussed a possible reaction forming primary ozonides, they did not include discussion of this particular O-atom rearrangement reaction. However, calculations by Cremer et al. (2011) on the O3 + acetylene system suggests a relatively high barrier to this rearrangement, so this ring closure reaction may be much slower than MechGen assumes. This will need to be re-examined when MechGen is updated."

The following was added to Section S1.8.2, concerning the reactions of CI's with non-alkyl substituents: "We assume that this type of O-shift reaction [$X_1C[OO]C(=O)X_2$ {ex} ↔ [$X_1*COOC(=O)(X)O*$] ↔ $X_1C[OO]C(=O)X_2$ {ex}] is faster than any competing possible reactions for excited carbonyl-substituted CIs, though theoretical data of Cremer et al. (2021) indicate that this assumption needs to be evaluated when MechGen is updated. In any case, this is not a sink for the CI …"

> *[SI] page 77 : "Note that the VHP reaction is not possible for tertiary substituents that do not have a hydrogens, so the excited radicals would only react via the hot acid reactions if possible, or only be stabilized if not, and the stabilized radical would not decompose even if it were in the syn configuration."*
> *All CI have access to the at least the dioxirane route (1,3-ring clousre), as it only involves the [COO] moiety; only the rate is affected by the substituents.*

The 1,3-ring closure reaction is included in Table 15 in the main manuscript. The following sentence was added to the subject location in the SI: "The only reactions assumed to be available to it [a stabilized CI that doesn't have other unimolecular reactions] are 1,3 cyclization forming a dioxirane or reaction with H_2O (see Table 15 in the main manuscript).

> *[SI] page 78: "One might expect that the extent of stabilization would increase with the size of the intermediate"*
> *This was studied also by Newland et al 2020*

Citation added. Sentence changed to "One might expect that the extent of stabilization would increase with the size of the intermediate (Newland et al, 2020), but the extent to which this is the case is uncertain."

Comments Regarding Aromatic Reactions

> *Line 786: "and the third involving a different type of cyclization where a peroxy O transfers to the opposite double bond, forming an epoxide and an alkoxy radical, with branching ratio fEpox." Recent literature indicates that this path is not important, and that epoxides are formed through epoxidation of beta-unsaturated alkoxy radicals. I have discussed this in detail in the past (e.g. Vereecken 2018, 2019 based on extensive theoretical literature and some experimental observations, and the non-contribution of this channel has been shown directly in recent experiments by Xu et al (2020). Postulating the direct formation of an epoxy-alkoxy radical skips intermediate steps for the epoxy-formation which are influenced by the reaction conditions (e.g. bimolecular reactions of the epoxy-RO2 intermediates after alkoxy epoxidation). Hence, attempting to shoehorn the observations onto this incorrect mechanism will make it hard in general to match the model to experimental data across a wider range of reaction conditions and for different aromatics, even when it appears to work for the limited observations where the parametrisation is based on. I understand that this representation of aromatic chemistry is common (e.g. Jenkin et al. 2018), but I strongly suggest adjusting the base reaction template for aromatic chemistry at some point to improve agreement with the available theoretical and experimental data (see e.g. Vereecken 2018 for an example mechanism).*

The treatment of the ring-opening reactions that do not result in formation of α-dicarbonyls or unsaturated 1,4-dicarbonyls is another major area of uncertainty that will need to be re-examined when MechGen is next updated. This will be a major effort not only because the actual mechanism is probably much more complicated, but because it will require readjusting portions of the aromatic mechanisms to fit available chamber data. Therefore, it is not practical to modify this treatment at the present time. As with the discussion of CI's, our response is to point out these difficulties and state that this needs to be updated. The modifications to the text are as follows.

The last half of 1st paragraph in Section 3.4 revised as follows: "O_2 reaction can occur either by abstracting the H-atom on the carbon where the OH added, forming HO_2 and a phenol, or addition to the ring forming an OH-aromatic-O_2 adduct. As discussed below, one of the possible reactions of this adduct is decomposition to form the same products as the abstraction reaction, so for simplicity we assume that this is the major pathway for formation of phenolic products, so only the addition reaction is considered."

The first part of the 4th paragraph in that section was modified as follows: "The subsequent reactions of the excited adducts are assumed to be as shown in Figure 1, which uses benzene as an example. This is based on the scheme recommended by Jenkin et al (2018b), and incorporated in the MCM (Bloss et al, 2005; MCM 2023)."

Figure 1 was modified to show the additional schemes suggested by Vereecken and referenced in the modified as shown below.

The following new paragraph was added to the discussion of the scheme on Figure 1: "The Phen and Bcyc pathways in aromatic mechanisms (routes [a] and [b] on Figure 1) are well supported by existing product data (Calvert et al, 2002) and theoretical calculations (e.g., Calvert et al, 2002; Vereecken, 2018, 2019). However, the existence of the Epox pathway (routes [c]), which is speculative and included in the mechanism to account for carbon balance and is not inconsistent with reactivities of aromatics observed in environmental chamber experiments (Carter 2010b, Carter and Heo, 2013) is much more uncertain. Although this pathway is assumed in the aromatics mechanisms recommended by Jenkin et al (2018b) and is incorporated in the MCM (Bloss et al, 2005; MCM 2023), theoretical calculations discussed by Vereecken (2018) indicate that the direct epoxide formation from aromatic-OH-O2 adducts (route [c]) is unlikely to be important. Vereecken (2018) suggests alternative pathways that are shown on the top and bottom right side of Figure 1 as alternatives that appear to be more consistent with theoretical calculations and recent experimental data (Xu et al, 2020). These alternatives will be considered when MechGen is updated, but currently MechGen uses the Epox as a placeholder to provide mass balance needed in aromatic mechanisms. The possible alkoxy cyclization reaction shown on the bottom right of the figure is discussed in Section 5.6." (Section 5.6 is new, see below).
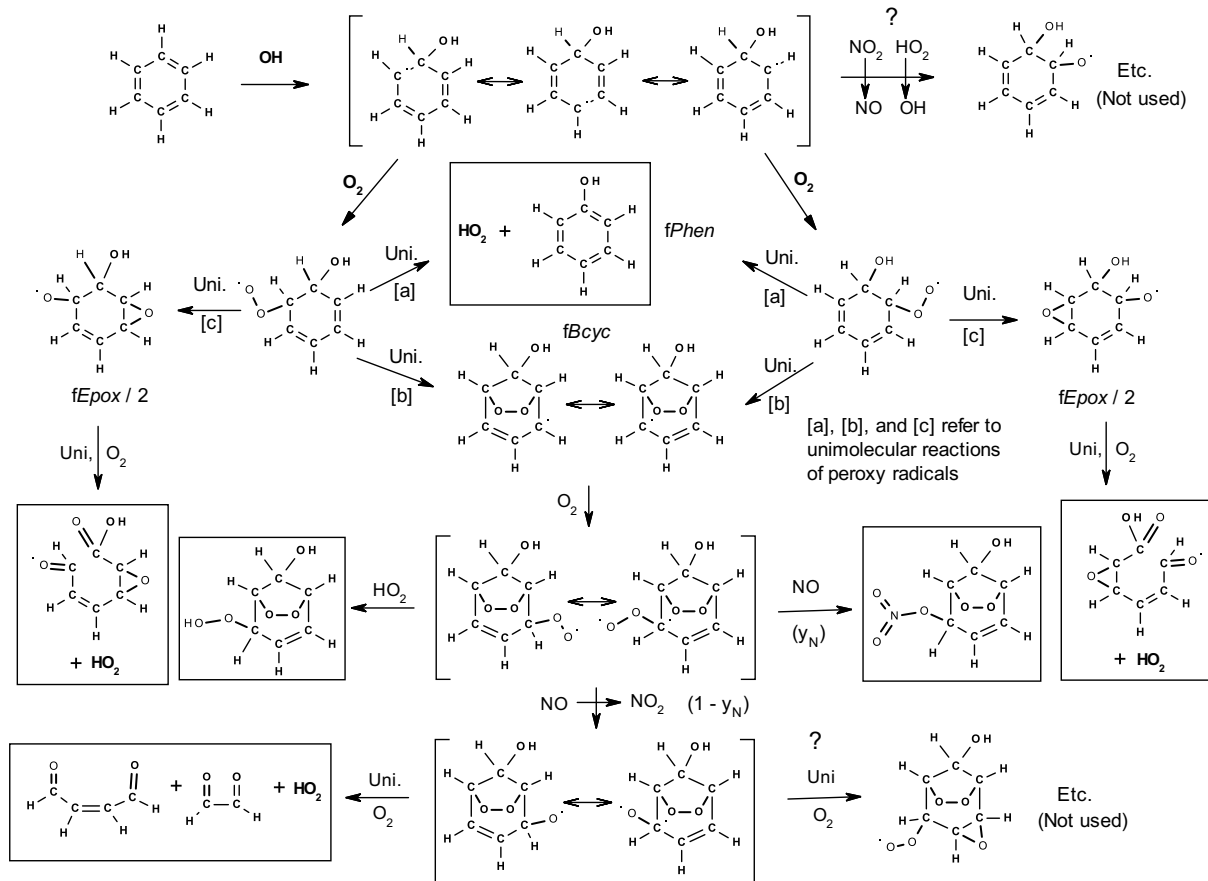
Figure 1. Initial reactions following the addition of OH to aromatic rings, using benzene as an example, showing the formation of the OH-aromatic adduct, followed by unimolecular reactions assumed for the OH-aromatic-O_2 adduct. Major products assumed to be formed are indicated, and additional pathways that may occur but are not currently used are also shown.
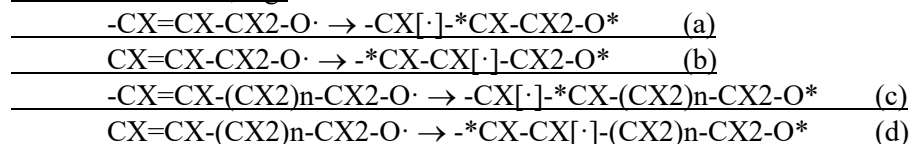
Section 5: "reactions of alkoxy radicals"
Epoxidation of beta-unsaturated alkoxy radicals could be an interesting addition. Theoretical work shows it is important in aromatic compounds (see above), and to some extent in the oxidation of multi-unsaturated species such as isoprene (e.g. Vereecken et al 2021, and Carlsson et al. 2023). It is an additional pathway for epoxide formation, in addition to the (higher barrier) beta-OOH alkyl radical decompositions (e.g. D'Ambro et al 2017, and IEPOX formation from isoprene)

This is also relevant to issues of aromatic mechanisms discussed above. A new section 5.6 was added under the discussion of alkoxy radicals, as follows:

"5.6.   Cyclization Reactions
"Unsaturated alkoxy radicals could possibly undergo cyclization reactions where the alkoxy radical adds to the double bond, e.g.

-CX=CX-CX2-O· → -CX[·]-*CX-CX2-O*	(a)
CX=CX-CX2-O· → -*CX-CX[·]-CX2-O*	(b)
-CX=CX-(CX2)n-CX2-O· → -CX[·]-*CX-(CX2)n-CX2-O*	(c)
CX=CX-(CX2)n-CX2-O· → -*CX-CX[·]-(CX2)n-CX2-O*	(d)

8

"These reactions are not considered in the current version of MechGen because they are not consistent with available product data on the reactions of OH with isoprene (Calvert et al., 2000), where the observation of methacrolein and methyl vinyl ketone as major products indicates that these reactions do not compete with the b-scission reaction forming a-hydroxy alkyl radicals. However, theoretical calculations and experimental data suggest that the epoxide formation in reaction of b-unsaturated alkoxy radicals [process (a), above] is important following the addition of NO3 radicals with isoprene; Carlsson et al. (2023) and Vereecken (2018) suggested that this reaction may also account for the missing mass balance in aromatic systems (see Section 3.4). This could be attributed to beta-scission decompositions forming a-nitrato or a-peroxy-substituted alkyl radicals are much faster than decompositions forming a-hydroxy radicals, but this is contrary to current estimates discussed in Section 5.2 and the SI. Note, however, that the competing cyclization reaction (b) is estimated to be somewhat more thermochemically favored, and cyclizations forming larger rings [e.g. (c) and (d), above] are even more thermochemically favored, though they have not been suggested in published mechanisms to our knowledge.

"None of these alkoxy cyclization reactions are currently considered by MechGen. However, based on recent results discussed above, they will need to be considered when MechGen is updated."

Comments Regarding Other Reactions

Line 647: "This is based on data of Slagle et al. (1984) for vinyl and Slagle et al. (1988) for methylvinyl and i-C4H9·. More recently, Matsugi and Miyoshi (2014) directly measured the yield of HCO radicals from vinyl"
Can the generalization of the data for small C2-C4 compounds to large molecules (e.g. C10) be relied upon? The rapid energy distribution over the entire molecule could lead to vinylperoxy stabilization. E.g. Sandia Nat. Lab. studied alkenylperoxy radicals made in this fashion (Meloni et al. 2007)

Added the sentence "This [decomposition] is assumed to be applicable to additions to radicals of any size, though it is possible that rapid energy distribution following reaction of larger vinyl species with O2 could lead to stabilized vinylperoxy radicals. However, insufficient data are currently available to quantitatively assess the degree to which this occurs, so MechGen assumes that the stabilization is not important."

Line 709-712: "Miyoshi (2011) reports results of theoretical calculations of hydroperoxyalkyl rate constants that extrapolate to much lower rate constants at atmospheric temperatures" Vereecken and Nozière 2020 already mentioned that at least some of the rate coefficients (OOQO2 H-migration) by Miyoshi 2011 may not be very accurate. Other theoretical work yields rate coefficients that extrapolate to higher values, competitive to O2 addition. E.g. Ali et al. 2023.

Added the following sentence to this paragraph: "On the other hand, Vereecken and Nozière (2020) indicated that some rate coefficients given by Miyoshi (2011) may not be accurate, and other theoretical work may yield higher rate constants when extrapolated to room temperature (e.g., Ali et al, 2023)."

Line 1109 "No attempt was made to make separate estimates of rate constants for H-shift isomerizations of cyclic peroxy radicals by either Vereecken and Nozière (2020) or in this work." See Vereecken et al. (2021) for systematic tabulated rates for H-shifts in cyclic peroxide-RO2 that might be useful to derive rates for cyc-RO2.

Added the sentence "Vereecken et al (2021) does include calculations for cyclic peroxy radicals that could be useful when MechGen is updated, but gives insufficient data to derive useful SARs at this time."

Section 4.7 "Ring closure of unsaturated peroxy radicals"
- The fast ring closure reactions of mono-unsaturated RO2 (see Vereecken et al., 2021) could be an interesting addition. These reactions are predicted to be quite fast; experimental data should become available soon. This reaction class is probably more common than conjugated alkadiene RO2s.
- The current literature does not support formation of epoxy-alkoxy radicals in cyclic nor linear conjugated peroxy radicals (see comment on aromatics)
- Epoxidation is likely due to cyclization in beta-unsaturated alkoxy radicals (see also comment on alkoxy radicals epoxidation)

The following two short paragraphs were added to the end of Section 4.7: "Note that, as discussed in Section 3.4, the epoxide forming cyclization reaction may not be as important in aromatic systems as currently estimated, and if this is the case then their analogues (reactions labeled [c] above), may also not occur in acyclic systems. This will need to be examined when MechGen is updated." (new paragraph) "It is also possible that ring closure involving only single double bonds may be non-negligible. However, if such reactions were to be significant, they should be important in the isoprene + OH reaction system, resulting in formation of different products than what are observed (Calvert et al, 2011, and references therein). Therefore, we currently assume that these cyclizations are only important in radicals with conjugated double bonds."

[SI] page 17: ". The product data for these reactions (Calvert et al., 2000) are consistent with addition to the primary radical occurring ~35% of the time. "
O2-addition on allylic radicals is reversible, allowing re-equilibration between the various RO2 (see Peeters et al. 2009, 2014, Novelli et al. 2020, Medeiros et al. 2022). Assigning a single site-specificity without accounting for the redissociation does not allow a correct description of the RO2 isomeric distribution across all reaction conditions as the actual distribution depends on the reaction conditions. E.g. at high NO the product distribution is determined by the nascent RO2 distribution, while conditions allowing long-lived RO2 will be governed by an equilibrium RO2 distribution, which may not be the same as the nascent distribution (see e.g. isoprene).

We are aware of the reversibility and the fact that assuming site-specificity is an approximation, but feel that this approximation is necessary given implementation issues and difficulty to derive SARs that take this into account. Using this simplification gives predictions that are consistent with available experimental data, and using a more complex model may not necessarily be any more accurate because of difficulties in estimating dissociation rate constants.

The following paragraph was added to Section S1.4.2: "It should be noted that this treatment [assuming site-specificity] is an approximation because O2 additions to double bonds are believed to be reversible (e.g., Peeters et al, 2014), allowing re-equilibrium between the various peroxy radicals that could be formed. However, it is difficult to derive general estimation rules that take this into account, and chemical mechanisms with rapid equilibrium processes can be difficult to implement for practical modeling because of mathematical "stiffness". It is also difficult to derive SARs for addition reactions that take into account this reversibility. Making the simplification assumption of site-specificity without re-dissociation gives predictions that are consistent with available experimental data, and using a more complex model with highly uncertain dissociation rate constants may not give predictions that are less uncertain."

[SI] page 20, Table A-9, entry for allylic radicals: the given ratio k(isom)/k(r+O2) seems not to take into account that the O2 addition on such radicals is reversible, and the "effective O2 addition rate" is thus also affected by the loss processes accessible after O2 addition.

Added the following to footnote [b] of Table S-9: "Note that this treatment is approximate because it does not take into account the possible reversibility of this reaction in the case of allylic radicals."

> *[SI] page 27: "peroxy + peroxy reactions are generally not important loss processes"*
> *This depends on the application. If SOA/aerosol formation is being modeled, dimer formation from RO2+RO2 is important even if it is a low-yield channel.*

Added the following sentences to the first paragraph of Section S1.5.4: "They [peroxy+peroxy reactions] may have an impact on SOA formation if dimer formation occurred, but generally these reactions cannot compete effectively with reactions with NO, NO3, and HO2 except in laboratory experiments carried out at very high concentrations," and at the end "Note that MechGen currently assumes that dimer formation does not occur, in part because it is not able to process cross reactions of different organic radicals."

> *[SI] page 47: "so that was taken as the actual lower limit for Ea0, and was used whenever the estimated values were lower than that."*
> *This prohibits estimating relative rates between H-shift reactions within the same molecule, if more than 1 channel is "fast".*

Changed the last sentence of the subject paragraph to: "so the exact value of these high rate constants, and therefore the upper limit Ea0 value used, would not have a practical effect on predictions of atmospheric chemical mechanisms unless there is more than one competing "fast" process. In that case, rate constants of competing fast processes are assumed to have the same rate constant."

> *[Si] page 48: "Note that strain energies should or might be significantly different in reactions of cyclic peroxy radicals"*
> *Vereecken et al. 2021 has rate coefficients for cycloperoxide-peroxy radicals. Irrespective of the presence of the oxygens in the ring backbone, the ring strain seems to prohibits significant H-migration in most of the configurations examined. Maybe this information could be of use, though that paper also makes the point that transference between the cyclic and acyclic cases is less than obvious.*

Added a citation to Vereecken et al. (2021) with regard to cyclic radicals. Changed the subject sentence to: "Note that strain energies should or might be significantly higher in reactions of cyclic peroxy radicals (Vereecken et al., 2021) or those where there is a double bond in the trans configuration in the transition state."

> *[SI] page 82: "but Vereecken et al. (2017) apparently did not consider subsequent reactions of this excited species."*
> *Vereecken et al. 2017 refers to the work of Kuwata et al. 2005, 2008, who studied this in detail. Since then there is also the experimental/theoretical work by Vansco et al. 2020, though these authors seem to ignore the epoxide route in favor of the H-migration route.*

Removed the comment about Vereecken et al. (2017) not considering this reaction.

> *Section 2.1.5 O^3P addition.*
> *Mention explicitly that addition on the central carbons of a conjugated system is assumed to be negligible, as opposed to cumulenes.*

This is not applicable because O^3P additions to double bonds of C_{3+} compounds are assumed not to form radicals. The O^3P is thought of as adding to both sides of the bond in effectively one process. The first paragraph of Section 2.1.5 was modified to clarify at the start that reactions of O^3P with C_{3+} compounds

are assumed not to form radicals because of simulations of chamber data. This was mentioned later in the section, but was moved up so this is clear when looking at the reaction scheme at the start of this section.

Comments Regarding Technical and Notation Issues

[SI] page 7, notation for chemical species
This naming scheme allows for multiple different notations for the same molecule, similar to some other line notation (e.g. SMILES). This is flexible and works well with graph/connectivity implementations. However, for databases and search implementations it is useful to have a canonical notation that provides a well-described unique way to write a molecule, irrespective of how the structure is perceived. Is there a canonization algorithm ?

MechGen has and uses an algorithm to create canonical structure names. A following paragraph is added to the end of Section S1.1 to make this clearer: "As with the SMILES notation, in many cases there can be more than one valid structure designation, as indicated on Table S-2 for many of the compounds. However, MechGen uses an algorithm to derive a single, canonical notation that is unique for each structure. This is necessary for the system to determine whether two structures, derived in different ways (either user input or as a result of a generated reaction) refer to the same compound. The first MechGen structure given on the right column of Table S-2 is the canonical MechGen designation; the others are valid designations that can be used to create new compounds. However, once created, MechGen uses the canonical structure for it."

Line 8: "O³P"
This is typically written as O(^{3}P); perhaps its useful to use the classical notation instead of the abbreviation in the abstract.

We think that both notations are acceptable, and chose not to make this change for consistency with other MechGen documents.

Table 2, bottom, and a few other places: Criegee intermediates are not biradicals

The term "biradical" was changed to "intermediate" when referring to Criegee intermediates. However, in other places the term "biradical" is used correctly (e.g., in the context of photolysis of cyclic compounds breaking bonds.)

Table 3:
- Not all notations used here can be found again in the supporting information (e.g. FHring(site) is missing in the caption of table A-42. It would be easier to link with the SI if all symbols here were present in the SI (most are).
- HCnbrs and nHCnbrs are not defined. I assume they mean hydrocarbon-neighbours and non-hydrocarbon-neighbours ?

The notations were added to the SI where needed and also Table 3 now indicates the tables in the SI where the parameter values are given. A discussion of F^{struct} for additions, which is already included in Table 3 was added to Table S-39 for completeness.

- It is not clear what the non-neighbouring "beta-substituents" are in the case of a cycloaddition like forO3. Are O3P and O3 swapped?

We are not sure we understand this concern or the question about O_3 and O^3P being swapped. I did clarify on Table 3 that nβ is the number of substituents on non-H groups bonded to groups around the unsaturated bond".

> *page 109, caption Table A-42*
> *This should explicitly mention the notation used in table 3 in the main paper. I assume this is "FHring(site)". Please check whether all notations from table 3 are reused in the SI.*

As indicated above, changes were made to assure that the notations used on Table 3 in the main paper are also used where applicable in the SI. The captions were modified to indicate that the ring corrections are FH^{ring}.

> *Line 548: "which would not be expected if this reaction did not occur".*
> *This requires some references to recent literature, or an explanation*

Added a citation to the theoretical study of Xu and Wang (2013) who reported calculations supporting this statement.

> *Section 3.2.1: "We expect that there should not be large entropy or energy barriers to these reactions, so MechGen assumes they all..."*
> *This is an example of where literature is available showing this expectation to be realistic and documented.*

Modified the subject sentence as follows: "We expect that there should not be large entropy or energy barriers to these reactions (e.g., Green et al., 1990; Vereecken et al., 2004, Vereecken, 2008), so …".

> *Line 1448: "and on the theoretical calculations of Vereecken et al. (2017) on the possible unimolecular reactions of the intermediates"*
> *A recent update is available for oxygenated and unsaturated SCI (Vereecken et al. 2022)*

Added the following text in this paragraph" "Briefly, the mechanisms adopted for MechGen are based largely on the IUPAC (2023) reviews of the reactions of O3 with ethene and alkenes with methyl substituents, and on the theoretical calculations of Vereecken et al. (2017) on the possible unimolecular reactions of the intermediates. Note that Vereecken et al (2022) provided a more recent update on the reactions of oxygenated and unsaturated CI's, but these have not yet been taken into account in the current estimates. Table 15 lists ...

> *[SI] page 18: cyclization of peroxy-substituted alkyl radicals*
> *Recent theoretical work shows that the barrier height does not decrease smoothly with decreasing ring size (e.g. Ali and Saswathy, 2023)*

Added the following to the end of the first paragraph in Section S1.4.3: "Note, however, that this treatment is an approximation because the effective activation energy may not change smoothly with ring size."

> *[SI] page 37: acylperoxy radicals*
> *A more detailed set of corrections for acylperoxy radicals was recently given by Seal et al. 2023*

Added to the following to the end of the paragraph discussing H-shift reactions of acyl peroxy radicals: "However, Seal et al. (2023) recently carried out a theoretical study of H-shift reactions in acyl peroxy

reactions whose results were not considered when these SARs were developed. These will need to be examined when MechGen is updated."

> *[SI] page 39: "However, the three-parameter expression in Equation (III) is not well suited for general estimation purposes" Perhaps mention that this is due to the strong cross-correlation between the three parameters when fitting*

Changed the referenced sentence as follows: "However, the three-parameter expression in Equation (III) is not well suited for general estimation purposes, because it requires more parameters to fit the data, with strong cross-correlation among the parameters, and the parameters are not straightforward to estimate. Therefore,, for this work …"

> *[SI] page 41: "relative to aliphatic unsubstituted radicals, which we presume refers to reactions of -C[OO·]< radicals"*
> *That is correct. See also Seal et al 2023 for more direct data and correction factors as a function of H-shift span.*

No change required.

> *[SI] page 52: "against the barrier heights calculated by Vereecken and Peeters (2009)"*
> *There is an update for this SAR in Novelli et al. 2021*

Added the following sentence to the end of the paragraph discussing the use of the Vereecken and Peeters (2009) barrier height estimates: "Note that Novelli et al (2021) have updated data relevant to these estimates that will need to be taken into account when MechGen is updated."

> *[SI] page 53, figure A-14:*
> *This deviation is expected: the activation energy Ea contains both energetic factors (barrier height Eb) and temperature-dependent entropic factors. Deriving Ea solely from Eb is missing this aspect.*

Changed the sentence discussing this as follows: "The calculated barrier heights are not exactly the same as the activation energies, and differ by as much as 4 kcal/mole at the lower barrier heights, but differences are expected because Ea takes into account temperature-dependent entropic factors while Eb does not. Nevertheless, the data are well fit by ..."

Editorial Comments and Error Corrections

The following suggestions were accepted or errors corrected:

> *Line 208, CI are represented as G(OO*)-O-G(O*): but there should be no central ether oxygen*
> *Section 2.1.5 O3P addition*

> *line 328 to 330: "The reaction with OH radicals ... were non-negligible"*
> *This sentence might perhaps fit better on line 326, after "... do not form radicals".*

> *line 482: "in direct sunlight"*
> *For consistency, use "direct overhead sunlight"*

> *line 510 and other places: "unpublished results"*
> *Perhaps change to "Carter et al., unpublished results"*

> *line 523: "and the rate constants are in cm3 molec-1 s-1"*

14

These are unimolecular rates, remove "cm3 molec-1"

Line 524: "R = 1.99×10-3 deg K mole kcal-1"
Kelvin is not "degrees Kelvin" (as opposed to e.g. degrees Celsius)

Line 539: "cyclohexadieneones" -> "cyclohexadienones"

Line 1090: "reactions faster the upper limit" -> add "than"

Line 1094: "groups by in alkyl peroxy radicals" -> remove "by"

Line 1515: "where an H bonded to the Criegee group" -> Perhaps say "an H-atom bonded to" for easier parsing (reads like H-bonded)?

[SI] Page numbers are missing on some pages

[SI] page 17: "the non-oh OH substituted position" -> remove "oh "

[SI] page 31: "unspeciated nitrate yields from formed in reactions " -> remove either "from" or "formed in"

[SI] page 37: "This is the either ratio" -> swap "the" and "either"

[SI] page 41, caption table A-17, last line: "were derived" is very widely spaced.

[SI] Page 56: "to the ketone group ketone products formed." -> correct sentence

[SI] Page 73 to 83: I'm not overly fond of using "Criegees" in formal text. If a short notation is needed, "CI" might be preferred.

[SI] Page 77: "excited CH3C[OO]" -> "excited CH3CH[OO]"

[SI] page 77 : "do not have a hydrogens" -> "have a hydrogen" or "have hydrogens"

[SI] page 83: "the CIs.On the" -> missing space