Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-1164-RC2, 2020

Atmospheric Chemistry and Physics
Discussions

Interactive comment on "H-migration in peroxy radicals under atmospheric conditions" *by* Luc Vereecken and Barbara Nozière

Anonymous Referee #3

Received and published: 20 March 2020

This RO2 H-shift SAR is a long-awaited advance in atmospheric chemistry. The authors quite sensibly choose to simplify the enormously complex RO2 chemistry by striving to create a (rather comprehensive) "single-functionality" SAR, which can then (assuming some simple additivity rules) be used as a good initial guess also for more complex multifunctional systems (and e.g. R/S stereochemistry, cyclic structures inside the TS ring, etc). Instead of attempting to generate comprehensive conformer sets for all their structures, they also really heavily on what they call "rel-MC-TST", where the "full" set of conformers (at a given level of quantum chemical theory) is generated only for a reference reaction, and more complex systems are studied by only computing a subset of conformers for those systems, including both ones as similar as possible to the reference system, and others "representing the change in molecular structure

between reference and target systems". While I'm dubious as to the universal applicability of this approach, e.g. for systems with multiple (non-alkyl) functional groups and thus a large number of H-bonding patterns, as in heavily oxidised large RO2, it is likely a good and cost-effective choice for the systems with relatively few non-alkyl functional groups studied here.

The authors are also to be commended for providing their raw data as a supplement - though I'm not sure if the provided data actually contains conformer data as well? (Given how important the authors consider the conformers to be, I would argue that the dataset perhaps SHOULD contain also at least the most important conformers for each system; see also comment 12 above.)

I have a number of somewhat technical issues that I would like to hear the author's comments on. Some of these may require minor revisions to the manuscript, for others I'd just like to hear the authors' response here in the discussion area.

1)Please define exactly what is meant by a "first-order SAR"?

2)Please explain what is meant by the sentence "for the remaining reactions the product energy is assumed to be similar to analogous reactions with similar product characteristics"? I fist assumed this refers to cases where the H-shift leads to prompt decomposition, i.e. where a QOOH "product" minimum does not exist. But upon closer reading, I realised this probably part of the authors' general "rel-MC-TST" philosophy, where computational time is saved by not explicitly computing product energies for all reactions… Anyway, this could if possible be reformulated to make the issue even clearer.

3)The authors argue against H-shifts in "trans"-alkenes by stating "To reduce the strain in the transition states very large migrations spans would then be required, leading again to low rate coefficients due to entropic disadvantage. ". But wouldn't the presence of the double bond remove at least one of the internal rotors, thus reducing the entropic disadvantage? Wouldn't it be worth checking e.g. some 1,8 or 1,9 H-shifts for trans-alkenes (with otherwise favourable substituents for H-shifts) to make sure the argument holds? Note that similar arguments also apply to conjugated or triple bonds, as well as cyclic substituents within the TS ring, as in many monoterpene-derived RO2: the steric strain likely prevents most H-shifts, but on the other hand the entropic penalty is lesser, potentially making for example even 1,9 H-shifts feasible (e.g. in the d3-carene + O3 system). On the other hand, it's perfectly understandable that the authors wish to limit the scope of their SAR - the number of potential RO2 systems is after all enormous - so the trans-conformers can well be left to a future study.

4)Illustrative (stick diagram) pictures of the different H-shift cases studied could aid readers especially from the “physics-side” of the ACP community. I'm a chemist, but even I needed some time to understand exactly what “allylic H-migration with a endocyclic gem-substituted double bond“ refers to.

5)The authors report large discrepancies (of unknown origin) between different studies concerning rates of “scrambling” H-shifts between RO2 and ROOH groups. In our experience, transition states for such reactions are particularly prone to artefacts in the Hartree-Fock stage of coupled-cluster calculations, where the HF algorithms used in programs such as Gaussian, Molpro etc. converge to an incorrect, higher-energy solution, analogous to local minima in geometry optimisations. That can then lead to either higher - or sometimes lower, but in any case incorrect - coupled cluster energies. Whether or not this error occurs seems to be almost random, with different HF algorithms (i.e. either different programs, or different algorithm choices within the same program), and different basis sets leading to different results - but sadly none of the combinations we have tested so far (including using the otherwise very robust MCSCF algorithm in Molpro for the initial HF guess) seems to be guaranteed to always give the lowest possible HF energy. For some discussion on this, see for example the Moeller et al 2019 paper cited in the manuscript. One quick way to check if this is an issue (apart from systematically performing “orbital rotation” checks on all HF energies as suggested for example by Vaucher and Reiher, J. Chem. Theory Comput. 2017, 13, 3,

1219-1228) would be to compare the “DFT-only” level results from different papers on this reaction class. Generally, we would expect (much) larger scatter in the DFT results due to larger inherent errors, and to differences between different functionals, but since DFT seems to be generally less affected than coupled cluster by this “local-minimum orbital” problem, then this case might be the exception. . .

6)Is there a particular reason the authors choose to use canonical CCSD(T), rather than the potentially more cost-effective variants such as CCSD(T)-F12, DLPNO-CCSD(T), or the very recent combination of the two, DLPNO-CCSD(T)-F12? I am fully aware that the details of the CCSD(T) correction is unlikely to be the major error source on the rates, I'm just interested in their reasoning here.

7)One more recent reference containing some “medium-level accuracy” computed RO2 H-shift rates for multifunctional systems (e.g. -ONO2 and C=O; also a C3 ring inside the TS) is: D. Draper et al., Formation of Highly Oxidized Molecules from NO3 Radical Initiated Oxidation of Δ-3-Carene: A Mechanistic Study, ACS Earth Space Chem. 2019, 3, 8, 1460-1470. The authors might wish to compare their SAR predictions also to the rates given therein.

8)In their explanation of rel-MC-TST (supplementary material), I believe they slightly mischaracterise the approach proposed by Moeller et al 2016 (and related approaches): they do not “generate a limited number of conformers at a very low level of theory rather than attempting a near-exhaustive characterization of the full population”, instead they explicitly DO attempt to generate the full population at a very low level of theory, and then attempt to come up with cost-effective approaches of limiting the number of calculations needed at higher levels. The authors may well be correct that with currently available method combination s, this then fails to “reliably generate the most contributing conformer” - but that is a different issue. Also, I would note that while, as said, the authors' present approach is likely very appropriate for their present systems, then as chemical complexity increases, something like the Moeller 2016 approach will inevitably be needed. Consider for example the case of a RO2 from a

C10 monoterpene containing 8 or so O atoms - corresponding to the least condensing "monomer" products in monoterpene oxidation. This could easily (depending of course on the number of remaining cyclic structures) contain on the order of 15 or so relevant dihedral angles, implying at least millions of relevant conformers. No matter how clever the selection of "reference" systems, at some point the "judicious selection of conformers by hand" just becomes impossible, and a systematic and automated approach - inevitably relying on analysing a large number of conformers at a lower level, chosen to maximise error cancellation with the desired final level - will be required.

9)The authors could briefly explain what procedure they carry out to ensure they have indeed found the "full" conformer population for their reference systems in rel-MC-TST: presumably this involves scanning over all dihedral angles (in 120-degree increments, or perhaps even smaller?) and then optimising all the obtained combinations (excluding obviously impossible ones) with the desired final DFT method? And then they must have some approach for distinguishing between (and eliminating) duplicates, as multiple inputs can lead to the same DFT local minimum, but with small differences due to the convergence criteria?

10)Line 340: comapred is a typo.

11)While I don't dispute that the SAR performs well, I wonder slightly at how impressive the factor of 2 agreement quoted in the abstract is, given that the total number of experimental data points included by the authors appears to be 22, and the SAR contains a fairly large number of parameters. (I understand the authors do not actually claim it is particularly impressive on its own, I would just like to hear their thoughts on this.)

12)Do the authors have some guidelines on what to do for more complex systems not covered directly by the SAR? Obviously, the SAR prediction is a very good starting point and sanity check - if it's very low, then likely the H-shift will not be relevant anyway, but what about the cases where the SAR prediction is fairly high, and we need a more

precise value? Would the next step be to find the closest reference system covered by the SAR, and attempt to apply the authors rel-MC-TST approach? (If yes, then it would DEFINITELY be good if the authors uploaded also their conformer data - as said I'm not sure from a quick glance if that's already included in the supplementary material).

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-1164, 2020.