https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

Comparison of the GECKO-A and SAPRC MechGen Atmospheric Chemical Mechanism Generators

Authors: Julia Lee-Taylor[1], William P. L. Carter[2], Bernard Aumont[3], John J. Orlando[1], Zhizhao Wang[2], Marie Camredon[3], Richard Valorso[3] & Kelley C. Barsanti[1]

[1] Atmospheric Chemistry Observations and Modeling laboratory, NSF National Center for Atmospheric Research, Boulder, CO, 80307, United States

[2] Center for Environmental Research & Technology, University of California Riverside, CA 92521, United States

[3] Univ Paris Est Creteil and Université Paris Cité, CNRS, LISA, F-94010 Créteil, France

Corresponding author: Julia Lee-Taylor, julial@ucar.edu

ABSTRACT

Atmospheric chemical mechanisms describe the photolysis and oxidation reactions of volatile organic compounds (VOCs), including bimolecular and unimolecular reactions of product radicals. Fully explicit representation of a single VOC can require thousands to millions of chemical reactions and species. Consequently, explicit chemical mechanism generation has proceeded through the use of automated tools that rely on observational data and structure activity relationships (SARs) to estimate rate coefficients and branching ratios. Here we compare, for the first time, two automated mechanism generators, GECKO-A (Generator of Explicit Chemical Kinetics of Organics in the Atmosphere and SAPRC MechGen (SAPRC Mechanism Generator). We specifically compare mechanisms for ten representative VOCs and the resultant properties of generated product mixtures as relevant for atmospheric chemistry. The observed differences in product mixtures can be explained by systematic differences between the mechanism generators after the initial reaction steps. GECKO-A has a more detailed representation of photolysis, including Norrish type-2 photolysis, which results in the formation of reactive alkenes from aldehydes. MechGen includes more alkoxy radical isomerization processes and autoxidation via H-shift reactions, the latter of which are not currently included in GECKO-A. Also MechGen assumes fast cyclization of alkyl radicals to form lactones or carbonates, which is not assumed by GECKO-A, and for which available theoretical predictions are inconsistent. These differences highlight needs for additional experimental data, particularly for uncertain products and yields, as well as for near continuous updating of these mechanism generation tools as sufficient data and new SARs become available.

INTRODUCTION

Volatile organic compounds (VOCs) are emitted into the Earth's atmosphere from a variety of anthropogenic and natural sources, and provide the fuel that drives the hydrogen and nitrogen oxides (HO_x/NO_x)–catalyzed chemistry of the troposphere. This chemistry accomplishes the oxidative removal of these VOCs, but also leads to the production of secondary pollutants, including ozone and secondary organic aerosol (SOA), with concomitant impacts on air quality, ecosystem health and climate. This oxidative chemistry is extremely complex, owing to the large variety of VOCs emitted, the diversity of chemical regimes in which the chemistry occurs, and the number of generations of chemical reactions required to convert a single VOC precursor to its inevitable end-products (such as carbon dioxide, CO_2 and water, H_2O). Since a fully explicit representation of even a single species' oxidation may require thousands or even millions of chemical reactions, it is impractical for these mechanisms to be written by hand. Thus, software tools have been developed to automate the generation of these explicit mechanisms.

Two of the more common computer-based automated mechanism generator tools used in atmospheric chemistry applications are GECKO-A (Generator of Explicit Chemical Kinetics of Organics in the Atmosphere [Aumont *et al.,* 2005, 2026]) and MechGen (Carter *et al.,* 2025a,b). These tools are conceptually similar. They use known kinetic data as a starting point and incorporate structure-activity relationships (SARs) into the code to estimate rate

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

coefficients and branching ratios for the remaining reactions. For example, SARs are employed to estimate rate coefficients and branching for reactions of common oxidants (OH, O_3, NO_3) with VOCs, for VOC photolysis reactions, for reactions of peroxy radicals (RO_2) with various partners (e.g., NO, HO_2 RO_2), and for unimolecular reactions of alkoxy radicals (RO), among many others.

Despite this conceptual similarity, GECKO-A and MechGen are developed independently. Thus, the generator systems differ in the types of precursor VOCs that can be handled, the types of reactions included, specific choices of SARs implemented, and level of detail incorporated at various stages in the oxidation. As a result, the mechanisms generated by each system, and hence the outcome of the oxidation of a particular precursor, contain some commonality but also some quantitative differences.

In this work, we conduct a first set of comparisons between the two systems. Mechanisms are generated using both systems and then employed independently in a box model under a standardized set of conditions (roughly corresponding to a rural continental setting). We perform simulations for each of ten representative precursor molecules, including saturated, unsaturated and oxygen-containing species that originate from anthropogenic, biogenic and pyrogenic sources. For each species considered, we compare bulk properties (e.g., overall OH reactivity, vapor pressure, solubility) and, when significant, we investigate specific differences in the chemistry of the two systems.

1. METHODOLOGY

2.1 Mechanism Generation

Both GECKO-A and MechGen operate by specifying the structure of an initial reactant and applying the rules of the respective generator system to find the gas-phase reactions considered to be important for that reactant under atmospheric conditions, and to determine which organic oxidation products are formed. The SARs are then applied to the products themselves, iteratively. Experimental or theoretically calculated rate coefficients and product assignments are used when available, but in most cases SARs are used to estimate rate constants and products and determine which (if any) reactions can be neglected. The result is a list of chemical products and of all the reactions they are estimated to undergo. These reactions are used as inputs to a box model simulation program that calculates concentrations of the reactant and products for a specified set of conditions to determine yields of various types of products formed under those conditions and how they develop over time. The box model results are used in this work to compare the chemical mechanism predictions of the two systems.

Table 1 summarizes the reactions generated by each system. While the two systems are similar in approach and include mostly the same types of reactions, their SARs and rate coefficient or branching ratio estimates were derived independently and in some cases use different literature sources. For relatively well-studied reactions such as those of organics with OH, O_3, and NO_3, or the major reactions of common intermediate product types, the assignments and estimates used are based on the same experimental data sets and give very similar predictions for most cases. Significant systematic differences do, however, exist with respect to a number of classes of reactions. For multifunctional compounds or intermediates for which data are scarce and assumptions have to be made, results can differ significantly. In particular, GECKO-A includes both more photolysis reference reactions and more classes of RO_2 than does MechGen. On the other hand, MechGen includes more detailed radical chemistry (RO_2 isomerization, RO isomerization, and carbon-centered radical cyclization) than does GECKO-A, which has a more limited set of RO isomerizations (1,5-shifts only), isomerizes RO_2 in only a few special cases and does not include carbon-centered radical cyclization.

Table 1. Types of reactions supported by the GECKO-A and MechGen mechanism generation systems and their literature sources or description of treatments in each system.

Reaction Type	GECKO-A	MechGen
Non-radicals		
VOCs +OH	Jenkin *et al.* (2018a)	Carter (2021)
Aromatic + OH ring-opening	Jenkin *et al.* (2018b)	Carter *et al.* (2025a)

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

VOCs + O_3	Jenkin *et al*. (2020)	Carter (2021), Carter *et al.* (2025a)
VOCs + NO_3	Kerdouci *et al*. (2010, 2014)	Carter (2021)
VOC Photolysis	Aumont *et al*. (2005); 126 reference chromophores	Carter et al (2025a); 20 types of reactions with 13 sets of action spectra
PANs unimolecular reactions	Wallington *et al.* (2021), Jenkin *et al.* (2019)	Carter *et al.* (2025a)
Radicals		
C-centered radicals + O_2	Assume fast unimolecular reaction or else fast O_2 addition	
C-centered radical decompositions	Rapid decomposition of α-OH, -HO_2, or -NO_2 substituents; Rapid decomposition of R-C(O)· radicals	
C-centered radical cyclizations	N/A	Fast if 5-member ring ester + OH, RO·, NO_2, or NO_3 formed
C-centered radical ring opening	Fast if the radical center is on an epoxide or dioxane ring	Fast if the radical center is on a cyclopropyl ring
Peroxy radical bimolecular	Jenkin *et al*. (2019); 8 RO_2 classes + RCO_3	Carter *et al.* (2025a); 1 RO_2 class + RCO_3
Peroxy radical cyclization	Jenkin *et al.* (2019) Aromatic species only	Carter *et al*. (2025a)
Peroxy radical H-shifts	N/A	Carter *et al.* (2025a), Vereecken & Nozière (2020)
Alkoxy radical + O_2	Vereecken & Peeters (2009), Atkinson (2007)	Carter *et al.* (2025a)
Alkoxy radical decompositions	Vereecken & Peeters (2010)	Carter *et al.* (2025a)
Alkoxy radical H-shifts	Vereecken & Peeters (2010). 1,5-H shift isomerizations only	Carter *et al*. (2025a); All non-negligible H-shift isomerizations
Alkoxy radical H-elimination	N/A	Carter *et al.* (2025a)
Alkoxy radical α-ester rearrangements	Atkinson (2007), Carter (2000)	Carter *et al.* (2025a)
Criegee Intermediate reactions	Newland *et al.* (2022)	Vereecken *et al.* (2017)

Both generator systems use cutoff criteria to determine which reactions can be neglected when multiple competing reactions of the same set of reactants are possible and to limit mechanism size. Each model applies the rate from neglected reactions to the accepted reaction channels, thus preserving overall reactivity at the point of competition. The details of the approach, however, differ between the two models.

GECKO-A uses a branching ratio threshold ("brcut") for any set of competing reactions (Aumont *et al.*, 2026), which is commonly universally applied but can also be varied by reaction type. In addition, a yield threshold ("yldcut") filters competing reactions based on their products' maximum yield from the initial precursor. At least one product is always retained, even if its yield < yldcut. GECKO-A's reaction filters are not based on specific oxidant levels so that generated mechanisms are applicable under a wide range of conditions. Finally, GECKO-A applies a user-specified limit to the number of chemical generations for which reactions are predicted. In this study we use brcut and yldcut values of 0.5% and 0.1%, respectively, with a 5-generation limit. The GECKO-A generator is available as detailed in Aumont et al. (2026). This study used a pre-publication version of GECKO-A, with only trivial differences from the standard version (e.g. input and output formatting). The generated mechanisms are available in full in a Zenodo archive (Lee-Taylor et al., 2026).

MechGen also uses a branching ratio threshold for competing reactions of the precursor species (Carter *et al,* 2025b). The branching ratio threshold, initially 0.5% in this study, is progressively increased for reactions of subsequent intermediates in inverse proportion to their estimated upper limit yields. The product yields are determined for a specified "environment" for the mechanism, that is, under a selected set of representative

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

concentrations of atmospheric reactants (e.g., O_3, HO_x and NO_x species, and total peroxy radical concentrations), and are derived by multiplying the total yield from all contributing reactions by the total yields of the immediate precursors. MechGen then applies a yield filter to exclude further reactions of any product whose estimated yield is <0.01% of the initial precursor amount, irrespective of chemical generation. The MechGen mechanisms in this study yield 99% or more of their individual product species within 5 generations in our simulations, except for camphene and isoprene (yielding 98% and 96% respectively).

In both generator systems, once a reaction chain has reached the specified limit of either generations or yield, the final stable species in the chain undergo loss reactions at SAR-based chemical loss rates. However, the products of these final loss reactions are left unresolved, and are instead transferred to the "lost carbon" aggregated pool. Thus, lost carbon in GECKO-A mechanisms represents all products formed after the specified maximum generation, while lost carbon in MechGen mechanisms represents all products with individual predicted yields of less than 0.01%. Larger precursors with more complex mechanisms can have so many low-yield products that their aggregated yield becomes significant, and the computational saving gained by their removal must be weighed against the loss of chemical information.

2.2 Precursor molecules

Ten example precursors were selected to cover a range of molecular complexity, a variety of functional groups and molecules from multiple sources. They include the alkanes propane, n-butane, n-octane, and cyclohexane; additional urban/anthropogenic species hexanal, trans-2-hexene (t-2-hexene), and 2-methyl furan; and the biogenic species isoprene, alpha-pinene, and camphene. Since the GECKO-A SARs have not been specifically tailored to isoprene chemistry, we simulate isoprene with two versions of GECKO-A: one incorporating a significant number of specified isoprene reactions from the Leeds Master Chemical Mechanism (MCM) version 3.3.1. (Jenkin et al., 2015) as recommended by Aumont *et al.* (2026), and one using only the GECKO-A SARs. We also discuss results of a sensitivity test for 2-methyl furan, since the value supplied with GECKO-A v0 for the reaction of maleic anhydride with OH is known to be high by at least an order of magnitude (Aumont et al., 2026), and is ~44 times greater than that used in MechGen.

2.3 Box model simulations

2.3.1 The GECKO-A box model

The box model used in this study was developed specifically for use with the extremely large mechanisms generated by GECKO-A, making it the natural choice as a platform for comparing GECKO-A and MechGen mechanisms. The model employs the "twostep" solver (Verwer 1994; Verwer *et al*., 1996) and may be configured to represent a wide variety of scenarios, from controlled chamber studies (e.g. La *et al.*, 2016) to outflow plumes mixing with the ambient environment (e.g. Lee-Taylor *et al.*, 2011, 2015; Mouchel-Vallon et al., 2020). For this study we operate in gas-phase mode, however we track explicitly the vapor pressure and solubility of the chemical products to enable aggregation of these properties at the post processing stage. The GECKO-A box model used here is an update of the version used by Mouchel-Vallon et al. (2020). It and the box model described by Aumont et al. (2026) have diverged from a common codebase, but continue to use most of the same physical-chemical principles. The major differences between the versions are in their user interfaces and in their varying presentations of output data. *(We plan to make a public release of the GECKO-A box model package used at NSF NCAR via GitHub and Zenodo during the time this manuscript is under review, and will update this code availability statement accordingly.)*

2.3.2. Mechanism translation

The two generator models produce ASCII-format chemical mechanism files that have model-specific syntax conventions. For example, MechGen explicitly writes all bonds in the molecular backbone (as 'dashes'), while GECKO-A only writes double bonds and leaves single bonds implicit. The models also each have their own protocols for establishing a unique direction in which to write any given chemical formula. To enable MechGen mechanisms to be deployed in the GECKO-A box model and post-processor we built a Fortran tool to read and parse the MechGen reaction and chemical formula lists, and rewrite them according to the GECKO-A conventions. This conversion program employs GECKO-A generator routines to standardize the representation of the chemical formulae, and also to assess in a uniform way intrinsic molecular properties such as mass, functional group

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

identities, vapor pressure, and water solubility (see Section 2.4). This tool is available as part of the code package associated with this paper.

2.3.3 Simulation conditions

The mechanism pairs were compared under a single set of conditions, chosen to be somewhat representative of continental conditions several hours downwind of a major urban area. These conditions were also used here as the "representative environment" for MechGen mechanism generation. We constrained the oxidant environment to allow us to focus on the chemical and physical properties of the product ensemble without introducing divergence caused by potential differences in the treatments of inorganic chemistry between two models. We also constrained bulk RO_2 and RCO_3 concentrations for the purpose of calculating rates of their cross-reactions. Here, we applied the bulk RO_2 constraint to the GECKO-A mechanisms by assessing the fractional contribution of each of the 8 RO_2 types and scaling it to the constrained bulk RO_2 concentration at each time step. Constrained reactant concentrations are listed in Table 2. These concentrations imply a ratio of ~80:20 for RO_2 + NO reaction rates vs. those of RO_2 + HO_2. The initial precursor concentration was 1 ppmv in every case.

The simulations considered only gas-phase chemistry, without phase transfer or wall losses, and at a temperature of 298K and 0% humidity (this last condition is a legacy of previous chamber study comparisons). The photolysis rate profile used by GECKO-A was generated for a set of 135 reference species by the TUV model (Tropospheric Ultraviolet-Visible; Madronich & Flocke, 1997), for overhead sun conditions. The photolysis rate profile used by MechGen was generated for 20 types of photolysis reactions using direct actinic fluxes for direct overhead sunlight as employed in the reactivity calculations of Carter (1994). Both approaches yield approximately the same j_{NO2}. Simulations lasted for six hours.

Table 2: fixed reactant conditions and initial precursor concentrations used for model comparisons

OH	1.38×10^{7} molec.cm^{-3}
O_3	140 ppbv
HO_2	37 pptv
NO	0.4 ppbv
NO_2	2.9 ppbv
NO_3	0.6 pptv
$\Sigma(RO_2)$	30 pptv
$\Sigma(RCO_3)$	5.2 pptv
j_{NO2}	1.2×10^{-2} s^{-1}
H_2O	0%
precursor	1 ppmv (initial)

2.4 Derivation of bulk or collective properties

In comparisons of gas-phase mechanisms between different models it is pertinent to consider whether the mechanisms have similar overall effects on the simulated chemical environment, including similar potential for gas-particle partitioning and formation of SOA. The general chemical and physical characteristics of a chemical product mixture should be preserved when constructing reduced mechanisms for use in regional models. Here, in addition to comparing predicted yields of various types of products in the box model simulation results, we assessed OH reactivity (the rate of OH consumption due to reactions with organic products), atomic and functional group ratios, vapor pressures, and solubilities of the organic product mixtures. These properties affect contributions to atmospheric oxidizing capacity and regional SOA mass and were determined as described below, with "organic products" defined as all carbon-containing species except CH_4, CO, CO_2, "lost carbon" and the precursor in each case. The differing treatments of minor reaction pathways between the two systems generally result in higher levels of "lost carbon" in MechGen than in GECKO-A. (See table S1).

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

OH reactivity ("OHR") for a species mixture is defined as the sum of all species concentrations multiplied by their OH rate constants kOH, i.e. for all species i, OHR = $\sum([i] \times k\mathrm{OH}_i)$. OHR varies as a function of time, with units of s^{-1}. To facilitate comparison between the mechanism generation systems, we define here total OHR ("TOHR") which includes the precursor reactant and all organic products, while product OHR ("POHR") excludes the reactivity of the precursor. Since "lost carbon" is also excluded from the assessment owing to its non-specific nature, the OH reactivities are, strictly speaking, lower limits.

The time-varying atomic ratios C:C_0 (the ratio of average molecular carbon number in the reaction products to the carbon number of the precursor) and O:C (the ratio in the organic product mixture of oxygen atoms to carbon atoms) are calculated at the post-processing stage by assessing the carbon and oxygen content of all organic product molecules at each output timestep. Similarly, we calculate the number of functional groups per carbon (g:C) in the output, for each functional group type, by summing the total number of functional groups of a given type over all organic product molecules and dividing by the total number of carbon atoms summed over those same molecules.

Vapor pressures (*Pvap*) are predicted for individual molecules using the group contribution method of Nannoolal *et al*. (2004, 2008). This method for estimating the temperature-dependent vapor pressures of organic molecules was trained on a database of ~1600 atmospherically-relevant compounds, and is supplemented in GECKO-A with values for hydroperoxide and peroxy acetylnitrate (PAN) groups from Compernolle *et al*. (2010). Vapor pressure distributions of product mixtures are derived for specific time points in the simulation output by defining representative intervals ("bins") in *Pvap*, and summing the concentrations of organic products whose *Pvap* values correspond to each bin. For most precursors we assess vapor pressure distributions at the end of each of 5 precursor lifetimes, however since the alkane lifetimes exceed 1 hour, we assess their distributions at hourly intervals.

Water solubilities are predicted for individual molecules using the "GROMHE" group contribution method of Raventos-Duran *et al*. (2010). This method for estimating effective Henry's law constants (H^*) in air-water systems at 298K was trained on a database of 488 atmospherically-relevant compounds. The method performs best for monofunctional species with H^* below 10^3 M atm^{-1}, and tends to over- and underestimate H^* for di- and multi-functional species respectively. Solubility distributions of product mixtures are derived in the same way and for the same specific points in the simulation output as for the *Pvap* distributions.

2. RESULTS

3.1 Bulk properties

In this section we present inter-model comparisons of bulk properties of the chemical product mixtures, to indicate where differences exist between the two chemical mechanism generation protocols. Properties assessed here are: OH reactivity, vapor pressure, solubility, oxygen-to-carbon ratio (O:C), average carbon number as a fraction of precursor carbon number (C:C_0), and functional group distribution (g:C). A summary of simulation-averaged bulk property values is given in Table S2.

3.1.1. OH reactivity

The two models show comparable evolution of total OH-reactivity (TOHR) for most assessed precursors (Figure 1). Where the precursor has relatively low OHR (here, the alkanes), TOHR initially increases with time as the product mixture gains functionality and reactivity, then decreases as the destruction rates of OH-active products outcompete their formation rates. For more highly reactive precursors (broadly, the unsaturated species), TOHR decreases roughly exponentially since the precursor is significantly more reactive than the product mixture which replaces it. Cyclohexane shows substantially different TOHR between the two systems, with GECKO-A giving values up to 1.5 times those obtained with MechGen. This reactivity difference indicates significant inter-model differences in the reaction pathways and product identities, and is addressed in more detail in Section 3.2.4. The top contributors to TOHR for each system are tabulated in the Supplemental Information (Tables S3b-S12b).

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

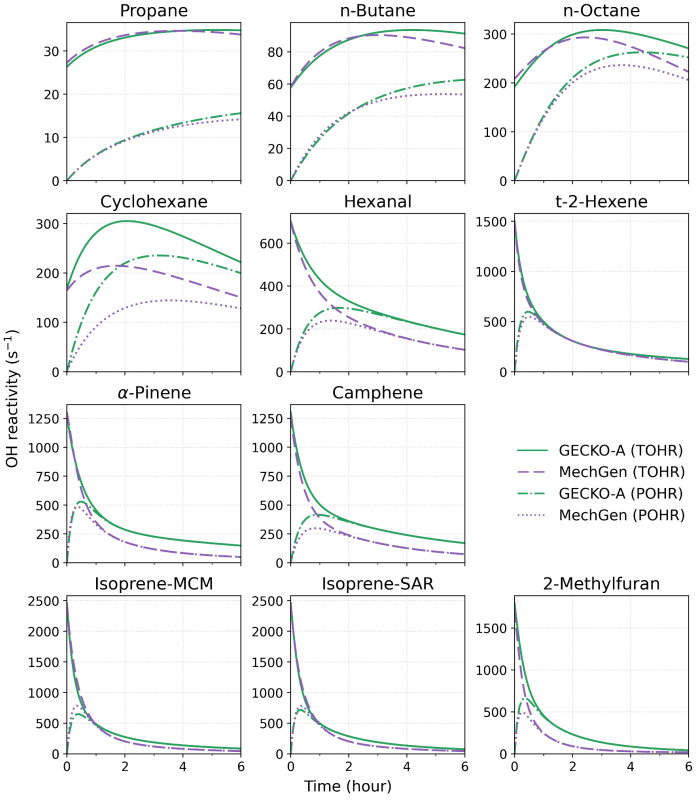


Figure 1. Total OH reactivity (TOHR) and product OH reactivity (POHR). Green lines, GECKO-A; purple lines, MechGen; upper lines in each panel, TOHR; lower lines, POHR. The gap between each TOHR and POHR pair represents the OHR of the remaining precursor.

For many species, since precursor OHR often overwhelms the product contribution, it is also interesting to compare product OHR ("POHR") in isolation. For this comparison we first remove biases arising from the models' different threshold treatments (see the discussion of "lost carbon", section 2.4) by considering molecular-mean POHR (*m*POHR) assessed over all products *j* of precursor *i*:

$$mPOHR(i) = POHR(i) / \mathbf{\Sigma}_{x=1}^{n} [j_x]$$

We then define the normalized POHR difference, "NDPOHR", as the inter-model difference of *m*POHR normalized to its inter-model mean, for each precursor and at each time (*t*):

$$NDPOHR(i,t) = (mPOHR(i,t)_{GECKO\text{-}A} - mPOHR(i,t)_{MechGen}) / ((mPOHR(i,t)_{GECKO\text{-}A} + mPOHR(i,t)_{MechGen})/2)$$

Figure 2 plots NDPOHR for each precursor, where positive (negative) values indicate that product mixture reactivity is higher (lower) in GECKO-A than in MechGen. The n-alkane and t-2-hexene product mixtures show good agreement, with *m*POHR from each model consistently within ± 10% of the mean. The other precursor systems show larger and mostly positive NDPOHR, varying widely over the 6-hour simulation. Isoprene begins with strongly negative NDPOHR (-40%) and swings to strong (>30%) positive values (beginning at ~ 2 hours or ~10 lifetimes) indicating that molecular-mean product OHR in MechGen peaks higher but decreases faster than that in either GECKO-A formulation. Camphene, hexanal, and α-pinene NDPOHR values increase progressively throughout (to 35%, 50%, and 70%, respectively), and cyclohexane NDPOHR decreases from ~50% to ~35%

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

during the 6-hour simulation (which covers ~2 cyclohexane lifetimes, Figure S6). 2-Methyl furan has by far the highest NDPOHR in its later oxidation stages (> 80% after 2 hours or ~10 lifetimes). Revising the OH + maleic anhydride rate constant progressively increases POHR (and thus TOHR) by a factor of up to ~ 2 by simulation end, with concomitant increases in NDPOHR (see Figure S6).

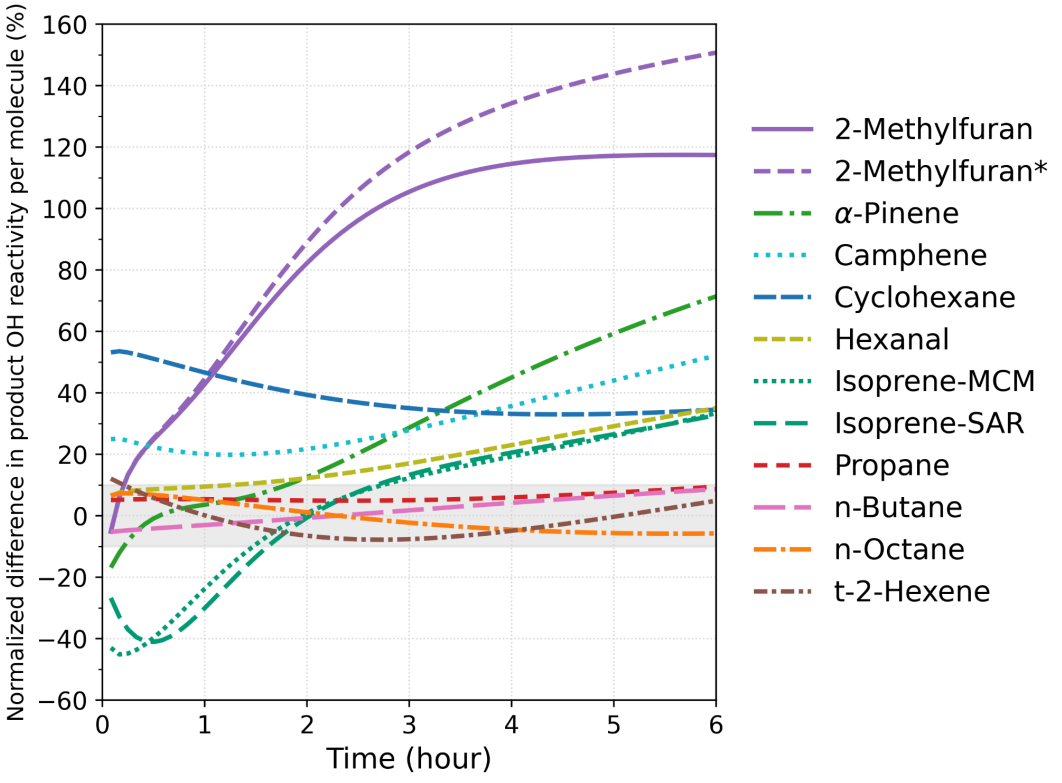


Figure 2. Time evolution of normalized difference in product OH reactivity ("NDPOHR"), defined as the inter-model POHR difference (GECKO-A - MechGen), normalized to the mean of the POHR values. Grey shading indicates agreement within ± 10%.

3.1.2. Vapor pressure and solubility

Vapor pressure (*Pvap*) and Henry's Law coefficient (*H**) distributions for the products formed from the various precursors are shown in Figures 3 and 4. In general, the two models produce similar distributions and values. For isoprene, the MechGen *Pvap* distribution skews noticeably more negative and *H** skews more positive, i.e. MechGen yields generally less-volatile and more-soluble products than does GECKO-A.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

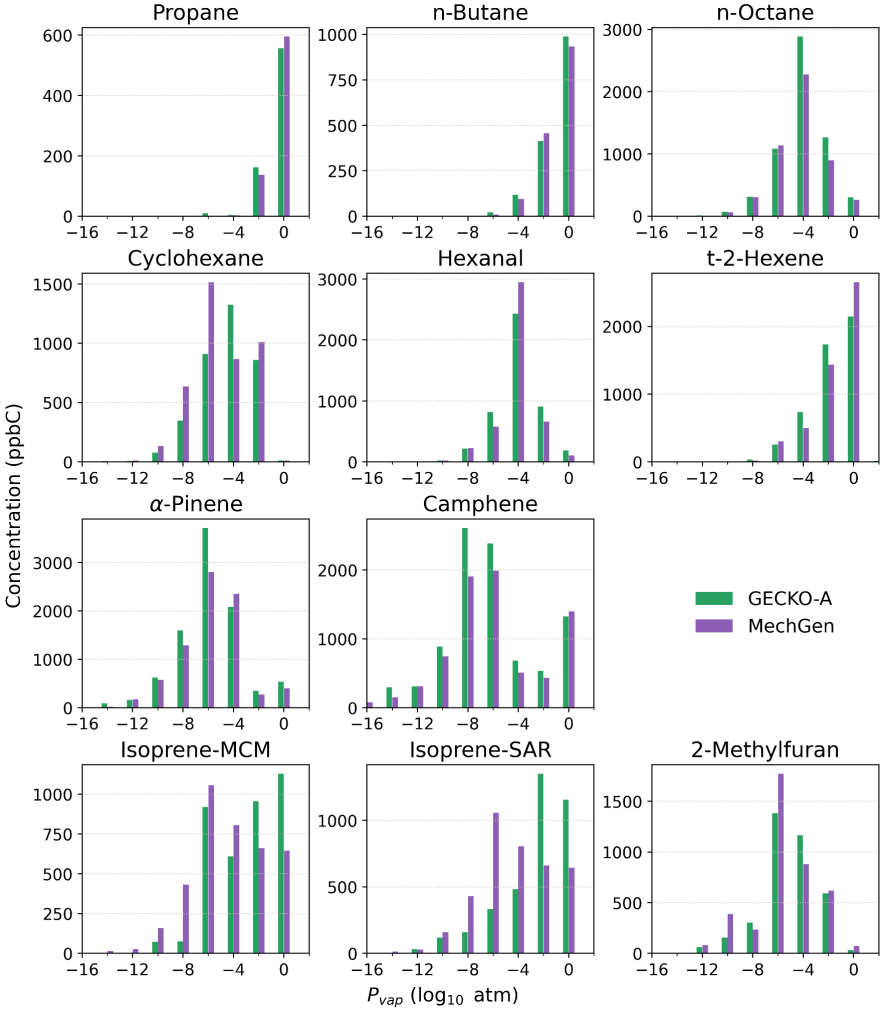


Figure 3. Comparison of vapor pressure distributions of multi-carbon products after 5 precursor lifetimes or, for the alkanes, at 6 hours. Green, GECKO-A; purple, MechGen. Pvap is expressed in units of $\log_{10}$ atm, concentrations in units of ppbC.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

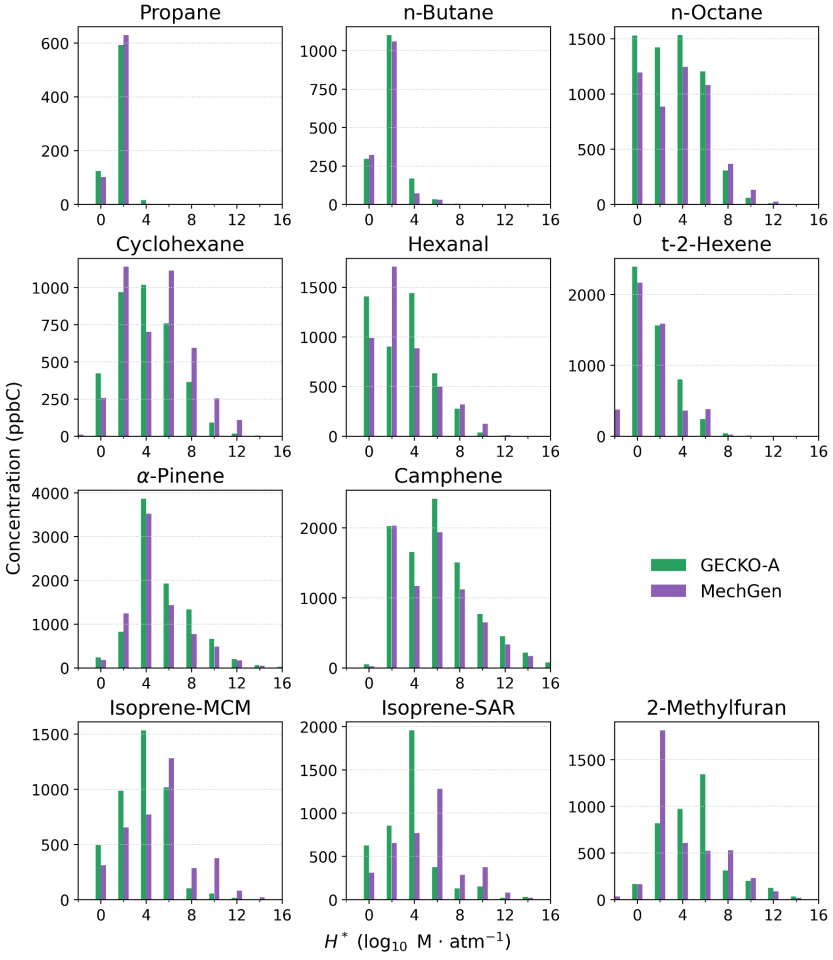


Figure 4. Comparison of Henry's Law coefficient (H*) distributions of multi-carbon products after 5 precursor lifetimes or, for the alkanes, at 6 hours. Green, GECKO-A; purple, MechGen. H* is expressed in units of $\log_{10}$ M.atm^{-1}, concentrations in units of ppbC.

The degree of agreement of estimated physical property distributions may be approximately quantified by comparing the carbon-weighted means of the distributions:

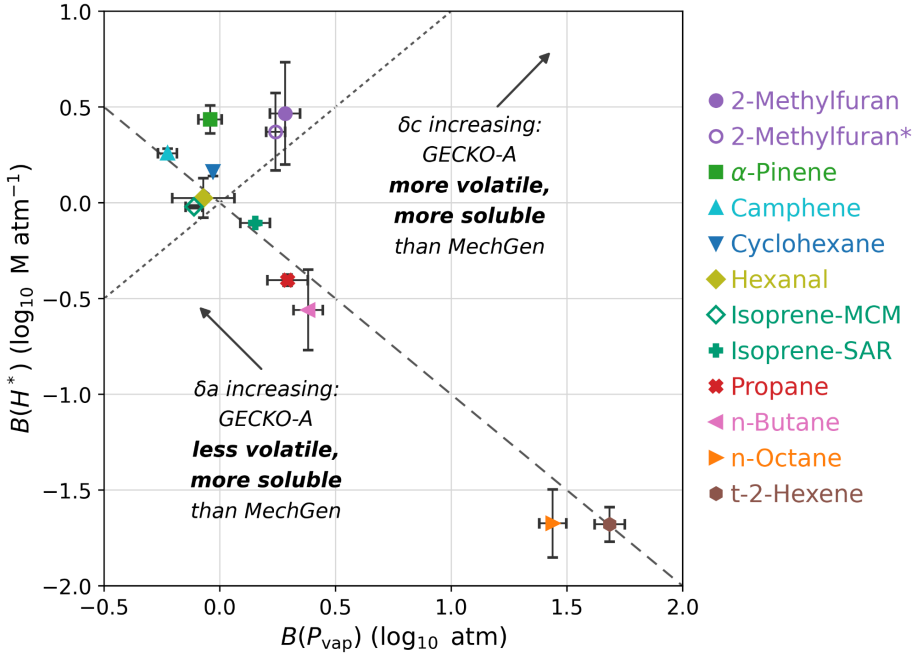
$$B(j) = log10\ [\mu(j)_{GECKO\text{-}A} / \mu(j)_{MechGen}]$$

where $B(j)$ = the base-10 logarithm of the inter-model ratio of the carbon-weighted means (μ) of the binned distributions of property j, where j is either $Pvap$ or H^*. Bins encompass two orders of magnitude. Means consider the discrete values at the ends of each of 5 lifetimes or for the alkanes, which are longer-lived, 6 simulation hours. A B value of 1 indicates that the carbon-weighted mean of the property distribution in GECKO-A is a factor of 10 (*i.e.* one order of magnitude) higher than that in MechGen, while a value of 0 indicates identical carbon-weighted means. In a correlation plot of $B(H^*)$ versus $B(Pvap)$ (Figure 5) most precursor systems in this comparison cluster around the origin, with B values within ±0.5 for one or both measures. Isoprene, however, shows significant inter-model differences, with $B(Pvap)$ and $B(H^*)$ of ~±1.5 for both the SAR- and MCM-based versions of GECKO-A.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

The parameters *H** and *Pvap* are somewhat inversely related, since increasing the oxidation degree of the carbon chain both increases water solubility and decreases vapor pressure, albeit with substantial scatter around the regression line influenced by specific functionalities (*e.g.* Hodzic et al, 2014; Lannuque et al., 2018; Isaacman-VanWertz et al., 2021; Aumont et al., 2026). The anticorrelation is roughly explained by the predominance of functionalization effects over fragmentation effects during oxidation. *B(Pvap)* and *B(H*)* are also generally anticorrelated (Figure 5), suggesting that differences in functionalization between the generated chemical schemes may dominate inter-model differences in fragmentation.

We introduce here two additional empirical quantities to help describe the inter-model differences encapsulated by *B(Pvap)* and *B(H*)*. δa and δc are the distances (in $\log_{10}$ units) of a data point in (Figure 5) from the 1:1 and 1:-1 correlation lines, respectively. δa roughly represents anticorrelated inter-model differences in physical properties arising from structurally similar products but with functional group differences. δc describes deviation from an inverse linear relationship, representing inter-model differences in the underlying carbon skeleton (arising from processes such as fragmentation or degree of cyclization), giving products whose estimated vapor pressure and solubility are uncorrelated. Figure 5 shows that most species fall in a narrow range within ±0.5 $\log_{10}$ units of the 1:1 line ($\delta a < 0.5$) and ±0.25 $\log_{10}$ units of the 1:-1 line ($\delta c < 0.25$), with n-octane, cyclohexane and α-pinene falling around the limits of this range. Isoprene's exceptionally large negative δa value (~-2.3) for both GECKO-SAR and GECKO-MCM reflect its very strong inter-model differences, *i.e.* GECKO-A predicts much higher volatility and lower solubility than MechGen in the isoprene system. The same is true, but to a far lesser extent, for n-octane and cyclohexane. By contrast, 2-methyl furan lies close to the 1:1 correlation line ($\delta a \approx 0.1$), but is farthest from the anticorrelation line ($\delta c \approx 0.5$), i.e. GECKO-A predicts higher values for both volatility and solubility than does MechGen for 2-methyl furan. This suggests that inter-model structural differences dominate in the 2-methyl furan system, while functionalization differences dominate the product physical property differences in the alkane and isoprene systems. The *B*-values for the 2-methyl furan sensitivity run fall well within the standard deviation bars of the original result, showing the GECKO-A *Pvap* and *H** values to be relatively insensitive to the anomalous OH reaction rate. α-Pinene occupies an intermediate position on Figure 5, with moderately high values for both δa and δc (both ~ 0.3) resulting from a near-zero *B(Pvap)* but significantly positive *B(H*)*, likely indicating differences in both functionalization and fragmentation.



https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

Figure 5. Correlation between inter-model exponent differences in vapor pressure and Henry's Law coefficient. Values are calculated as the average of the differences between discrete binned-distribution means at the end of each of five precursor lifetimes or, for alkanes, at six 1-hourly intervals; error bars represent the standard deviation of each average, irrespective of the widths of the property distributions. Positive (negative) values indicate that higher values are predicted by GECKO-A (MechGen). Short-dashed line, 1:1 correlation; long-dashed line, 1:-1 correlation.

3.1.3. Product atomic ratios

The time-varying ratios $C:C_0$ and O:C provide a more direct insight into the chemical structural tendencies of a mechanism. $C:C_0$ is the ratio of the mean carbon number of all organic products to that of the precursor molecule, illustrating the degree of fragmentation of a simulated organic mixture. O:C is the ratio of oxygen atoms to carbon atoms in the organic product mixture, a rough proxy for the degree of product functionalization. We define summary comparison metrics ND(O:C) and ND($C:C_0$) as the simulation means of the normalized inter-model differences for each of these atomic ratios. Figure 6 compares the atomic ratio results assessed for each model simulation pair. The two models generally yield very similar results, although species with outlying values for other metrics tend to also show greater disagreement in the product atomic ratios. For cyclohexane, hexanal, isoprene, and 2-methyl furan, GECKO-A produces lower $C:C_0$ values (more fragmentation) than does MechGen, at least partly because MechGen includes H-shift isomerizations of peroxy radicals that otherwise would react with NO to form alkoxy radicals that potentially decompose via beta-scission. Perhaps counterintuitively, GECKO-A also shows significantly higher O:C ratios than does MechGen. In MechGen, production of highly oxygenated molecules (HOMs) via H-shift isomerizations appears to be offset by greater ether and ester formation, while GECKO-A tends towards formation of oxygen-rich PAN and nitrate groups. This is further discussed in the next section and, for individual precursors, in section 3.2.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

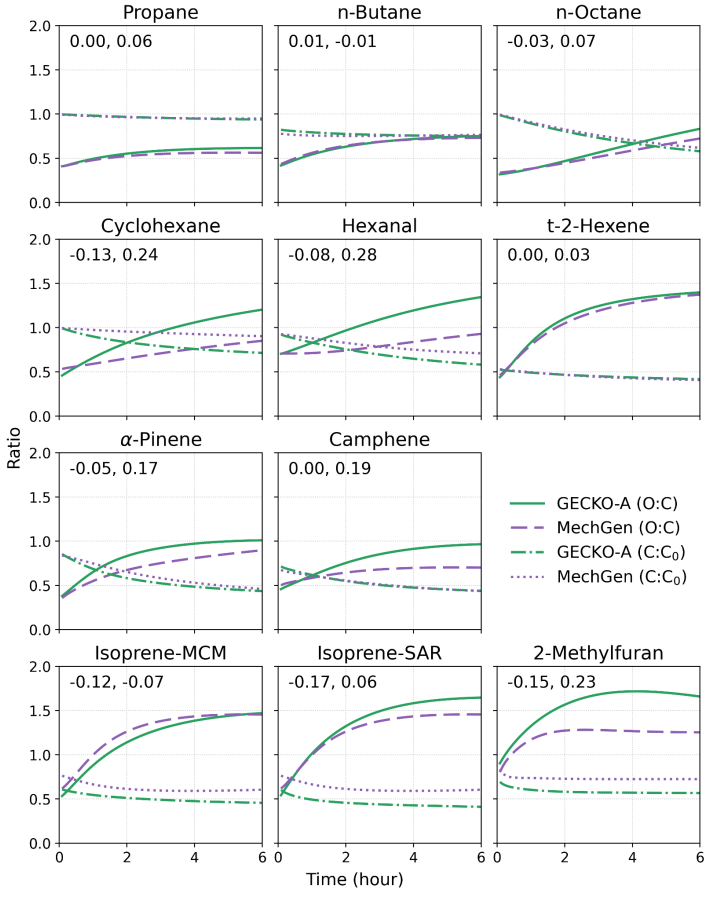


Figure 6. Inter-model comparison of O:C and C:C_0 ratios in organic product mixtures. Green lines, GECKO; purple lines, MechGen; solid and dash-dot lines, O:C; dashed and dotted lines, C:C_0. Notation on each plot gives the values of mND(C:C_0) and mND(O:C), in that order. The 2-Methyfuran sensitivity-case values for mND(C:C_0) and mND(O:C) are -0.12 and 0.11, respectively (see Figure S4).

3.1.4. Functional group distributions

To further explore chemical variations between the two models, we compare functional group distributions in the product mixtures. We define product functionality ("g:C") as the quantity of a particular functional group divided by the total number of product carbon atoms, averaged over the entire organic product mixture (excluding CO, CO_2, CH_4 and "lost carbon"). Figures 7 and S1 compare the distributions of 11 different functional groups and of aliphatic ring structures throughout the model simulations. There is generally good agreement between models for the smaller alkanes and trans-2-hexene, consistent with the close OHR correspondence. Agreement for other precursors is mixed. Several consistent themes emerge. 1) Larger or more complex precursors generally show more substitution, greater diversity of product functionality, and greater divergence between the two models. 2) Carbonyl groups (either aldehydes or ketones) are typically among the most abundant product groups for both models, as major products of RO decomposition following RO_2+NO reactions. 3) MechGen routinely produces more hydroperoxide groups (-OOH) and GECKO-A favors hydroxyl groups (-OH), as a result of the autoxidation included in MechGen. 4) MechGen frequently produces ester (R-O-CO-R) and ether (R-O-R) products (e.g. for cyclohexane, hexanal,

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

α-pinene, camphene, and 2-methyl furan), whereas GECKO-A typically does not. 5) GECKO-A often produces more PAN groups than MechGen.

The differences in functionality between the two mechanisms can directly affect the agreement between bulk physical and chemical properties. For example, substitution of a hydroxy group with a hydroperoxy group has limited impact on OH reactivity, however it can double the solubility and reduce volatility by an order of magnitude. Greater molecular differences have greater impacts. A test comparison of selected C6 molecules with nitrate and PAN groups (as favored by GECKO-A) versus C6 molecules with ether or ester groups (as favored by MechGen) shows that the former have higher O:C ratios (with 3 to 5 oxygen atoms versus 1 or 2), and lower vapor pressures by 1.5-2 orders of magnitude. The exception is ethyl-γ-lactone (the C6 cyclic ester), which is assigned a similar vapor pressure to the C6 PAN and nitrate. Solubility for these species increases over a range of 1.7 orders of magnitude in the order linear ether < nitrate < PAN < cyclic ether < linear ester < γ-lactone. The C6 nitrate, C6 PAN and C6 esters are assigned comparably low OH-reactivity by the mechanisms' SARs (4-6 × 10^{-12} s^{-1}), while the C6 ethers (linear or cyclic) are predicted to be roughly four times as OH-reactive.

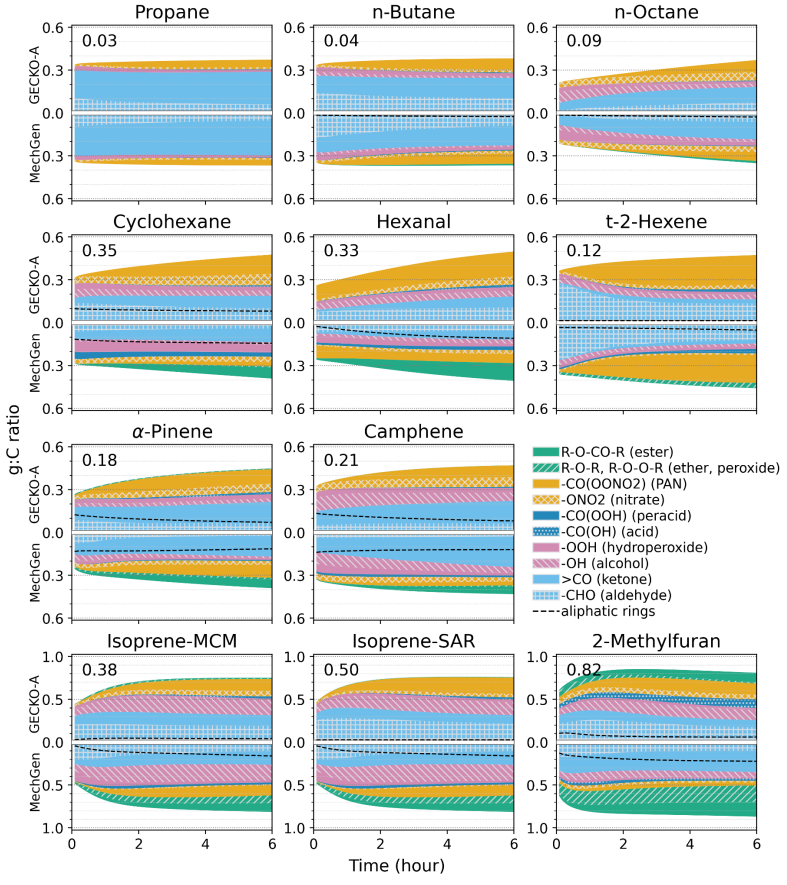


Figure 7. *Inter-model comparison of functional group distributions in product mixtures. Values represent the ratio of functional groups per carbon, averaged over the entire organic product mixture excluding CH_4 and "lost carbon". Aliphatic rings are included where their mean g:C exceeds 0.01. Nitro groups are negligible in these simulations. Σ\|Dfn\| values are shown in the top left corner of each plot. Σ\|Dfn\| for the 2-methyfuran sensitivity case is 0.53 (not shown). Note that the vertical scale for isoprene and 2-methyl furan is greater than for the other plots.*

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

The differences in functional group distributions between the two model systems can be summarized using the metric Σ|Dfn|, which we define as the sum over all functional groups of their absolute inter-model differences in g:C values. The definition of "functional groups" used here includes aliphatic ring structures and double bonds (see caption to Figure S1). Simulation-mean Σ|Dfn| ranges from 0.03 for propane to 0.82 for 2-methyl furan. Mechanistic reasons for the various differences will be explored for individual precursors in section 3.2.

3.1.5. Summary of comparison metrics

The inter-model comparisons of physical and chemical properties are summarized in Table 3 as simulation-mean metrics for each precursor (see also Table S2, Figure S2). To facilitate comparison, we define goodness-of-agreement ranges with labels from "poorest" to "very good". The range limits are defined a posteriori for each metric in proportion to the maximum absolute inter-model difference (D_{max}). Thus, they do not reflect model skill for either mechanism generator, but are instead intended as a (rough) guide to how well the results from the two mechanism generators compare. Ranges are defined as: "very good" (statistically insignificant), "good" (< $D_{max}/4$), "moderate" (< $D_{max}/2$), "poor" (< D_{max}), and "poorest" (= D_{max}), and assigned an integer label from 0 to 4, increasing as inter-model agreement worsens. We average these labels over the six first-order comparison metrics NDPOHR, *B(Pvap), B(H*),* ND(O:C), ND(C:C_0), and Σ|Dfn| to give a mean "agreement score". The agreement ranges and scores are detailed in Table S3. In general, goodness of agreement tends to be related, that is, a precursor tends to show similar range labels for several or most of its statistically significant metrics. This is to be expected, since molecular structure details (types and abundances of functional groups, presence or absence of rings, degree of fragmentation) directly influence reactivity, volatility and solubility.

We can group the precursors into loose groups based on their metrics values. The first group comprises the two smaller n-alkanes and t-2-hexene, precursors with good-to-very good agreement, and mean agreement scores of >1.0. Precursors in the second group, comprising n-octane, camphene, α-pinene, and hexanal, show mixed good-to-moderate agreement in almost all their simulation-integrated metrics, giving them overall agreement scores between 1 and 2. n-Octane shows "good" comparisons for all metrics except ND(O:C) (where it is "moderate"). NDPOHR is "good" for hexanal and "moderate" for the two terpenes. *B(Pvap)* and *B(H*)* are "good-to-very good" except for *B(H*)* in α-pinene ("moderate"). ND(C:C_0) is "very good" for camphene and "moderate" for α-pinene and hexanal. ND(O:C) is "moderate" for n-octane, "poor" for the terpenes, and "poorest" for hexanal. Net functionality agreement Σ|Dfn| is "good" for α-pinene and "moderate" for camphene and hexanal.

The final group, with the largest inter-model differences and agreement scores of >2.0, comprises cyclohexane, isoprene, and 2-methyl furan. Cyclohexane, isoprene-MCM and 2-methyl furan (sensitivity) show better mean agreement scores (2.2 - 2.3) than isoprene-SAR and 2-methyl furan (original), which have the worst-performing comparisons overall with mean scores of 2.8 and 2.7. Agreement improves in *B(Pvap)*, ND(C:C_0), and Σ|Dfn| for isoprene-MCM over isoprene-SAR, and in *B(Pvap)*, *B(H*)*, ND(O:C), and Σ|Dfn| for 2-methyl furan (sensitivity) over 2-methyl furan (original). Nonetheless, the extreme values for NDPOHR in both 2-methyl furan cases and for *B(Pvap)* and *B(H*)* in both isoprene cases mark these two species as outliers amongst our precursor set. Cyclohexane, while not an outlier by any single metric, is consistently in the "moderate-to-poor" agreement range (except for "good" *Pvap* agreement relative to the extreme isoprene values) indicating significant inter-model differences also for this species.

In summary, these comparison metrics suggest that the two mechanism generation systems GECKO-A and MechGen perform quite similarly for n-alkanes and t-2-hexene, but that application of the models' different SARs to more complex molecules can result in product mixtures with statistically significant - and sometimes dramatic - differences in chemical and physical properties.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

Table 3. Summary of simulation-mean inter-model comparison metrics[1]

[2]Precursor	[3]NDPOHR	[4]*B(Pvap)*	[4]*B(H*)*	[5]ND(O:C)	[5]ND(C:C_0)	[6]Σ|Dfn|	[7]score	[8]δa	[8]δc
n-Butane	0.04	-0.03	0.16	-0.01	0.01	0.04	0.7 (1,1,1,0,0,1)	0.14	0.10
Propane	0.06	-0.11	-0.02	0.06	0.00	0.03	0.8 (1,1,1,1,0,1)	0.06	-0.09
trans-2-Hexene	0.05	-0.23	0.26	0.03	0.00	0.12	0.8 (1,1,1,1,0,1)	0.34	0.02
n-Octane	0.04	0.29	-0.40	0.07	-0.03	0.09	1.2 (1,1,1,2,1,1)	-0.49	-0.08
Camphene	0.31	-0.07	0.03	0.19	0.00	0.21	1.2 (2,0,0,3,0,2)	0.07	-0.03
α-Pinene	0.32	-0.04	0.44	0.17	-0.05	0.19	1.7 (2,0,2,3,2,1)	0.34	0.28
Hexanal	0.19	0.15	-0.11	0.28	-0.08	0.33	1.8 (1,1,1,4,2,2)	-0.18	0.03
Cyclohexane	0.38	0.38	-0.56	0.24	-0.13	0.35	2.2 (2,1,2,3,3,2)	-0.67	-0.13
Isoprene-MCM	0.20	1.44	-1.67	-0.07	-0.12	0.39	2.2 (1,3,3,1,3,2)	-2.20	-0.17
2-Methyl furan (sensitivity)	1.03	0.24	0.37	0.11	-0.12	0.54	2.3 (4,1,1,2,3,3)	0.09	0.43
2-Methyl furan (original)	0.88	0.28	0.47	0.23	-0.15	0.82	2.7 (3,1,2,3,3,4)	0.13	0.53
Isoprene-SAR	0.21	1.68	-1.68	0.06	-0.17	0.51	2.8 (1,4,4,1,4,3)	-2.38	0.0

[1]Comparison metrics shown here are assessed as the averages of 6-hour difference timeseries, except where otherwise noted. Shaded values are statistically insignificant in the simulation mean.
[2]Precursors are listed in order of increasing (worsening) score.
[3]NDPOHR is the absolute inter-model difference in inter-model-mean-normalised OH reactivity per product molecule.
[4]*B(Pvap)* and *B(H*)* are the inter-model differences of the midpoints of the $\log_{10}$ distributions of *Pvap* and *H**, averaged at six 1-hour intervals (alkanes) or at five 1-lifetime intervals (all other precursors).
[5]ND(O:C) and ND(C:C_0) are, respectively, the inter-model difference normalised to the inter-model mean in the ratios of product oxygen-to-carbon and mean product carbon number to precursor carbon number.
[6]Σ|Dfn| is the sum over all functional groups of the individual absolute inter-model differences in functional groups per carbon.
[7]Mean agreement score is the average of the goodness-of-fit integer labels for the six simulation-mean metrics in columns 2 thru 7. Lower scores indicate better inter-model agreement.
[8]δa and δc are the perpendicular distances in log space from the 1:1 correlation and 1:-1 anticorrelation lines, respectively, on a plot of *B(Pvap)* vs *B(H*)*. Negative (positive) values represent points below (above) the reference line. δa and δc are excluded from the assessment of agreement score to avoid double-counting contributions from *Pvap* and *H**

3.2. Individual species comparisons

Here we compare details of the chemical mechanisms for each species, with a focus on explaining inter-model differences in the bulk properties of the product mixtures. We discuss the comparisons in approximate order of increasing molecular complexity and property divergence, from alkanes through alkenes and monoterpenes and concluding with 2-methyl furan. Our reaction scheme diagrams include "local yields", which we define as the simulation-specific yield of the denoted product from the previous species in the diagram for which a yield is shown, considering all reaction types. Thus, local yields are generally smaller than branching ratios (which sum to unity for each type of reaction of the immediate precursor) and larger than product yields (which express the production of any given species relative to the reacted amount of initial precursor, and thus potentially result from the product of multiple successive branching ratios). We express chemical formulae in Canonical SMILES format as generated by the Open Babel tool (O'Brien et al, 2011) implemented online at https://www.cheminfo.org/flavor/cheminformatics/Utility/FormatConverter/index.html, and last accessed June 2026.

3.2.1 Propane

Propane is the simplest precursor studied here, so we expect it to show no significant inter-model differences, and to demonstrate the consistency of our mechanism conversion and comparison protocols. Both models show the same 10 most abundant products (Table 4) and 10 highest contributors to OHR (*table S4a,b*), and the bulk parameters comparing the product mixture properties all indicate "good" or "very good" inter-model agreement (Table 3). The major product is acetone (at ~65% of simulation-integrated organic product molecules), with minor (2-10%)

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

contributions from peroxypropionyl nitrate (PPN), propanal, acetaldehyde, PAN, 2-propyl hydroperoxide, 2-propyl nitrate, and formaldehyde. Those inter-model differences that do occur are mainly driven by small differences in the rate constant of the initial OH reaction (specified as 4% faster in MechGen) and in primary-to-secondary propylperoxy formation branching ratios (specified as 0.31:0.69 and 0.28:0.72 in GECKO-A and MechGen, respectively).

Table 4. Top 10 propane products by simulation-integrated molecular abundance

Name or type	Canonical SMILES structure	Integrated abundance	
		GECKO-A	MechGen
acetone	CC(=O)C	62%	66%
PPN	CCC(=O)OON(=O)=O	9%	6%
propanal	CCC=O	8%	6%
acetaldehyde	CC=O	4%	5%
PAN	CC(=O)OON(=O)=O	4%	4%
2-propyl hydroperoxide	OOC(C)C	3%	3%
2-propyl nitrate	CC(ON(=O)=O)C	2%	3%
formaldehyde	C=O	3%	2%
n-propyl hydroperoxide	CCCOO	2%	1%
n-propyl nitrate	CCCON(=O)=O	1%	1%
Other products		2%	2%

3.2.2. n-Butane

n-Butane, as another small alkane, also has a relatively simple mechanism, although the additional C-atom introduces additional complexity, some of which is explored below. Both models produce butanone (methyl ethyl ketone, ~40% of organic product molecules), with lesser contributions from PAN (10-20%), acetaldehyde (14-22%), formaldehyde (6-8%), and 2-butyl nitrate (4-5%) (Table 5). All comparison metrics show “good” or “very good” inter-model agreement (Table 3).

Table 5. Top 10 n-butane products by simulation-integrated molecular abundance

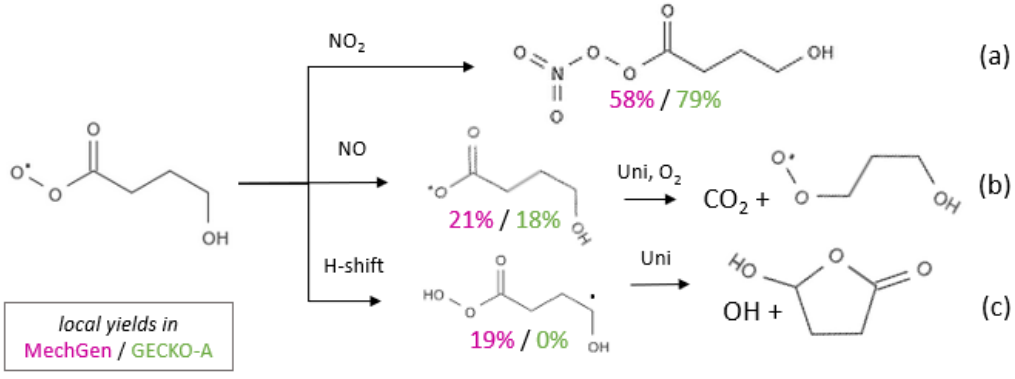
Name or type	Canonical SMILES structure	Integrated abundance	
		GECKO-A	MechGen
methyl ethyl ketone	CCC(=O)C	40%	37%
acetaldehyde	CC=O	15%	20%
PAN	CC(=O)OON(=O)=O	14%	17%
formaldehyde	C=O	9%	5%
2-butyl nitrate	CCC(ON(=O)=O)C	4%	5%
2-butyl hydroperoxide	CC(OO)CC	3%	4%
4-hydroxy butanal	OCCCC=O	1%	1%
ethyl hydroperoxide	CCOO	1%	1%
4-hydroxy butyl PAN	OCCCC(=O)OON(=O)=O	1%	( )
methyl hydroperoxide	COO	1%	( )
peroxyacetic Acid	CC(=O)OO	( )	1%
5-hydroxy-oxolan-2-one	OC1CCC(=O)O1	-	1%
Other products		11%	9%

( ) indicates species outside the top 10 in this simulation
- indicates species not produced in this mechanism

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

A close look at the n-butane mechanisms reveals potentially significant chemistry differences in the models' treatment of 2nd-generation acyl-peroxy radical reactions. Reaction scheme 1 shows the major reaction channels of 4-hydroxy-butyrylperoxy radical (accounting for 98% of the acyl peroxy transformation in each model) which is derived from 4-hydroxy-butanal, a major first-generation product of OH attack at n-butane's CH_3 groups.

First, the model-generated (SAR-derived) rate coefficient for acylperoxy radicals with NO_2 (reaction 1a) is significantly faster in GECKO-A than in MechGen ($1.2x10^{-11}$ vs $7.7x10^{-12}$ cm^3 $molec^{-1}$ s^{-1}). Second, MechGen allows for autoxidation reactions of the acylperoxy species to occur. Third, MechGen commonly cyclizes radicals of the general form ·RC(=O)-OX, if X is a suitable 'leaving group' (e.g., -ONO_2 or OH) and if a 5-membered ring can be formed (see table 1). In the case shown here, MechGen in part autoxidizes the 4-hydroxy-butyryl-peroxy radical to form O[C]CCC(=O)OO, which then cyclizes to form a C4 cyclic ester (γ-lactone) (reaction 1c). As a result of these two systematic chemistry differences, MechGen's simulation-integrated branching yields for products (1a), (1b) and (1c) are 58%, 21%, and 19% respectively, while GECKO-A gives branching yields of 79%, 18%, and zero.



Reaction scheme 1: n-Butane: major transformations of the 4-hydroxy-butyryl-peroxy radical.
Numbers are time-integrated local yields in the MechGen (magenta) and GECKO-A (green) box model simulations.

3.2.3. n-Octane

n-Octane, our third n-alkane, shows very similar chemistry between the models as before. Almost all metrics exhibit "good" inter-model agreement, while O:C agreement is "moderate". The most abundant products result from fragmentation (PAN, PPN, acetaldehyde and formaldehyde, each contributing 4-8% of product molecules, table 6). Various first-generation nitrates and hydroxy- or hydroperoxy-ketones are present at 2-4% abundance. There are, however, specific chemistry differences. The most significant difference in terms of bulk product functionality (Figure 7) is in the abundance of hydroperoxide (-OOH) groups in MechGen versus hydroxide (-OH) groups in GECKO-A, owing to the inclusion in MechGen of major autoxidation processes.

Table 6. Top 10 n-octane products by simulation-integrated molecular abundance

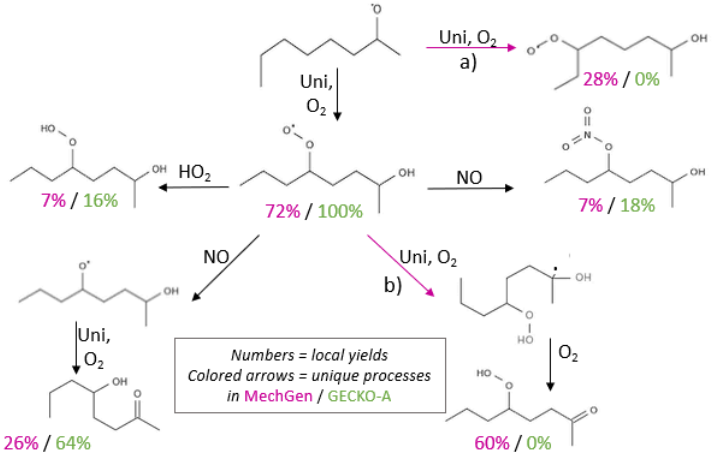
Name or type	Canonical SMILES structure	Integrated abundance	
		GECKO-A	MechGen
formaldehyde	C=O	6%	4%
PAN	CC(=O)OON(=O)=O	5%	8%
acetaldehyde	CC=O	5%	7%
PPN	CCC(=O)OON(=O)=O	5%	4%
3-octyl nitrate	CCCCCC(ON(=O)=O)CC	4%	3%
6-hydroxy-3-octanone	CCC(CCC(=O)CC)O	4%	( )
2,5-octanedione	CCCC(=O)CCC(=O)C	3%	4%

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

4-octyl nitrate	CCCCC(ON(=O)=O)CCC	3%	3%
7-hydroxy-4-octanone	CCCC(=O)CCC(O)C	3%	( )
7-hydroperoxy-3-octanone	CC(OO)CCC(=O)CCC	( )	3%
6-hydroperoxy-3-octanone	CCC(=O)CCC(OO)CC	( )	3%
3,6-octanedione	CCC(=O)CCC(=O)CC	( )	3%
propanal	CCC=O	2%	( )
Other products		59%	58%

( ) indicates species outside the top 10 in this simulation

An example of the differences between the generation systems is shown in Reaction Scheme 2. Following OH attack, both systems react the initial 2-peroxy radical with NO to generate an alkoxy species or a nitrate (with ~60% and ~20% effective branching ratio under the simulation conditions), or with HO_2 to produce smaller quantities of hydroperoxide. Key differences then ensue: In both models, the 2-octoxy radical isomerizes leading to the eventual formation of 1,4-hydroxy nitrates, -hydroxy carbonyls and -hydroxy nitrates, as would be 'traditionally' expected. MechGen however includes additional and competitive pathways, including a unimolecular 1,6-alkoxy shift (path 2a, top right) that leads to a series of 1,5-hydroxy nitrates, hydroxy carbonyls and hydroxy nitrates, and autoxidation of the 1,4-hydroxy-peroxy radical to produce the hydroperoxy carbonyl shown in path 2b.



Reaction scheme 2: n-Octane: First-generation transformations of the n-octane-derived 2-RO radical. Numbers are time-integrated local yields in the MechGen (magenta) and GECKO-A (green) box model simulations. Colored arrows represent processes unique to one model.

The chemistry differences shown here do not significantly impact the C:C_0, or OHR comparisons, but do contribute to the objectively slightly less good agreement in *Pvap* and *H**, where the product mixture is more volatile (*B(Pvap)* +0.3) and less soluble (*B(H*)* -0.4) in GECKO-A than in MechGen (Fig 5). For example, the estimated vapor pressure of MechGen's autoxidized C8 hydroperoxy ketone products is lower by a factor of ~10 than that of GECKO-A's C8 hydroxy ketones (4×10^{-6} vs 5×10^{-5} atm), and the estimated *H** is a factor of ~2 higher (3.2×10^5 vs 1.6×10^5 M atm^{-1}).

3.2.4. Cyclohexane

Cyclohexane exhibits larger inter-model differences than do the other alkanes, with "moderate" agreement indicated for POHR at 35-55% higher in GECKO-A than the inter-model mean (Fig 2), and "poor" agreement for C:C_0 and O:C which average 12% lower and 24% higher, respectively (Fig 6). The functional group distributions (Fig 7) show only "moderate" agreement and amplify the trends shown by n-octane: GECKO-A is more abundant in PAN groups, nitrates, aldehydes, and hydroxides while MechGen favors ketones, hydroperoxides, peracids, and esters (mainly

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

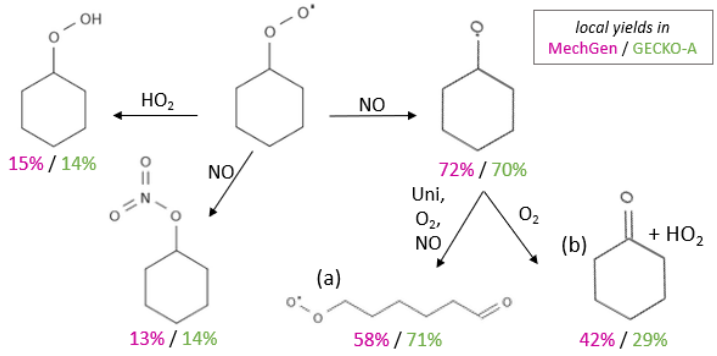
lactones, which are not produced by GECKO-A) (see Table 7). Agreements in volatility and solubility are poorer than those of n-octane: GECKO-A produces more-volatile (*B(Pvap)* +0.38) and less-soluble (*B(H*)* -0.56) products (Fig 5) than MechGen, while still maintaining the expected volatility-solubility anticorrelation (giving a statistically insignificant δc value).

Table 7. Top 10 cyclohexane products by simulation-integrated molecular abundance

Name or type	Canonical SMILES structure	Integrated abundance	
		GECKO-A	MechGen
cyclohexanone	O=C1CCCCC1	14%	20%
oxolan-2-one-5-peraceticacid	OOC(=O)CC1CCC(=O)O1	-	11%
formaldehyde	C=O	10%	( )
cyclohexyl nitrate	O=N(=O)OC1CCCCC1	9%	8%
cyclohexyl hydroperoxide	OOC1CCCCC1	3%	4%
C3 di-PAN	O=C(CC(=O)OON(=O)=O)OON(=O)=O	3%	( )
3-acylperoxynitrate-2-hydroxypropanal	O=CC(C(=O)OON(=O)=O)O	3%	( )
3-hydroperoxy-6-oxo-hexanoic peracid	O=CCCC(CC(=O)OO)OO	( )	3%
a C6 peracid-keto-PAN	OOCOCC(=O)CCC(=O)OON(=O)=O	( )	3%
5-hydroxy oxolan-2-one-5-peracetic acid	OC1(CC(=O)OO)CCC(=O)O1	-	3%
oxolan-2-one	O=C1CCCO1	-	3%
a C6 keto-di-PAN	O=C(OON(=O)=O)CCC(=O)CC(=O)OON(=O)=O	2%	( )
C4 di-PAN	O=C(OON(=O)=O)CCC(=O)OON(=O)=O	2%	-
3-hydroxy hexanedial	O=CCCC(CC=O)O	2%	-
6-nitrooxy hexanoic peracid	OOC(=O)CCCCCON(=O)=O	( )	2%
6-hydroperoxy hexanoic peracid	OOCCCCCC(=O)OO	( )	2%
2-nitrooxy cyclohexanone	O=C1CCCC(C1)OON(=O)=O	1%	( )
Other products		51%	42%

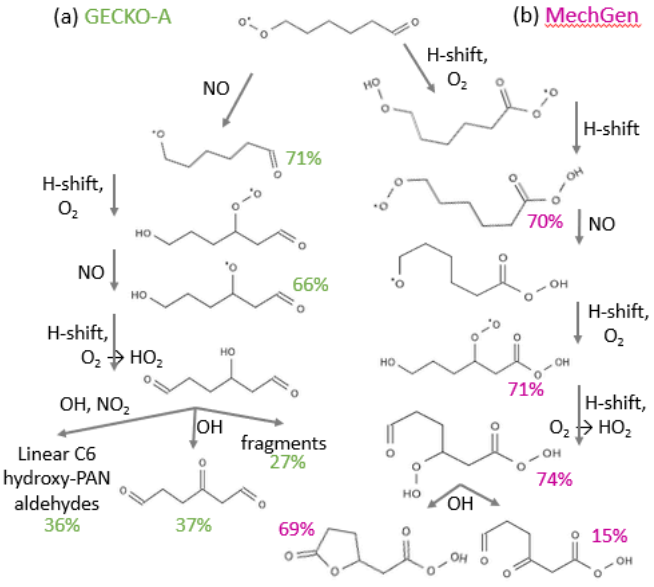
( ) indicates species outside the top 10 in this simulation
- indicates species not produced in this mechanism

The inter-model differences may be best understood by comparing their respective chemical assumptions. The cyclohexyl peroxy radical formed in the initial oxidation step reacts via four main channels, albeit with inter-model branching ratio differences. While the ring-opened 6-oxo-hexylperoxy radical product (product (a) in Reaction Scheme 3) is favored in both models, the ratio of its yield to that of the ring-retaining cyclohexanone (product (b)) is smaller in MechGen (4:3) than in GECKO-A (5:2). Cyclohexanone is only moderately reactive with OH; kOH = 6.4 × 10^{-12} cm^3 $molec^{-1}$ s^{-1} in both models. As a major first-generation product, it is the greatest contributor to POHR in the MechGen simulation but only the third-ranked in the GECKO-A simulation (Table S7b), reflecting inter-model differences in both its production rate and the fate of the competing 6-oxo-hexylperoxy radical.



https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

Reaction scheme 3: Cyclohexane: Major transformations of the cyclohexyl peroxy radical. Numbers are time-integrated local yields in the MechGen (magenta) and GECKO-A (green) box model simulations.



Reaction scheme 4: Cyclohexane: major transformations of the 6-oxo-hexylperoxy radical. Numbers are time-integrated local yields in the MechGen (magenta) and GECKO-A (green) box model simulations.

Reaction scheme 4 shows the major transformations of the 6-oxo-hexylperoxy radical, with striking differences between the two models. GECKO-A reacts the peroxy species with NO followed by alkoxy isomerization (path 4a) yielding 3-hydroxy-hexan-dialdehyde and its OH-oxidation product 3-oxo-hexan-dialdehyde. These species are strong POHR contributors in GECKO-A (Table S7b), with kOH values of 6.5×10^{-11} and 4.8×10^{-11} cm^3 molec^{-1} s^{-1}. MechGen instead specifies autoxidation of the 6-oxo-hexylperoxy radical (path 4b) to yield OH-reactive 3-hydroperoxy-6-oxo-hexanoic peracid (kOH = 3.5×10^{-11} cm^3 molec^{-1} s^{-1}), whose major product is a high-yield but relatively unreactive (and therefore persistent) substituted lactone (kOH = 3.3×10^{-12} cm^3 molec^{-1} s^{-1}), produced in MechGen by re-cyclization.

In summary, the autoxidation and lactone formation processes present in MechGen but not in GECKO-A are responsible for generating higher-OHR products from cyclohexane in GECKO-A than in MechGen, and also for the differences in functional group distributions (Figure 7) and degrees of fragmentation between the two models which in turn increase disagreement in *Pvap*, *H**, O:C, and C:C_0.

3.2.5. Hexanal

Hexanal shows mixed chemical property agreement between the two models. Agreement is "good" for POHR, *Pvap* and *H**, and "moderate" for functionality and C:C_0, but "poorest" for O:C. Major products common to both models include hexanoyl peroxynitrate ("C6 PAN") and formaldehyde, with pentyl nitrate, PAN, and acetaldehyde also relatively abundant (table 8). As in the case of cyclohexane, functional group distributions (Figs. 7, S1) show MechGen uniquely producing considerable amounts of lactones and exceeding GECKO-A in abundance of hydroperoxides and peroxy-acids, while GECKO-A favors PAN-groups, aldehydes, hydroxides and nitrates.

Table 8. Top 10 hexanal products by simulation-integrated molecular abundance

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

Name or type	Canonical SMILES structure	Integrated abundance	
		GECKO-A	MechGen
C6 PAN	CCCCCC(=O)OON(=O)=O	18%	13%
5-methyl oxolan-2-one	CC1CCC(=O)O1	-	16%
formaldehyde	C=O	10%	3%
5-ethyl oxolan-2-one	CCC1CCC(=O)O1	-	6%
PAN	CC(=O)OON(=O)=O	5%	3%
acetaldehyde	CC=O	4%	2%
4-hydroperoxy pentanal	CC(OO)CCC=O	( )	4%
1-nitrooxy pentane	CCCCCON(=O)=O	( )	4%
5-methyl 5-hydroxy oxolan-2-one	O=C1CCC(O1)(C)O	-	4%
a C5 keto PAN	CC(=O)CCC(=O)OON(=O)=O	3%	( )
4-hydroxy pentanal	O=CCCC(O)C	2%	( )
a C5 hydroxy PAN	CC(CCC(=O)OON(=O)=O)O	2%	( )
PPN	CCC(=O)OON(=O)=O	2%	( )
C3 di-PAN	O=C(CC(=O)OON(=O)=O)OON(=O)=O	2%	( )
a C4 aldehydic PAN	O=CCCC(=O)OON(=O)=O	2%	( )
1-hydroperoxy pentane	CCCCCOO	( )	2%
Other products		49%	44%

( ) indicates species outside the top 10 in this simulation
- indicates species not produced in this mechanism

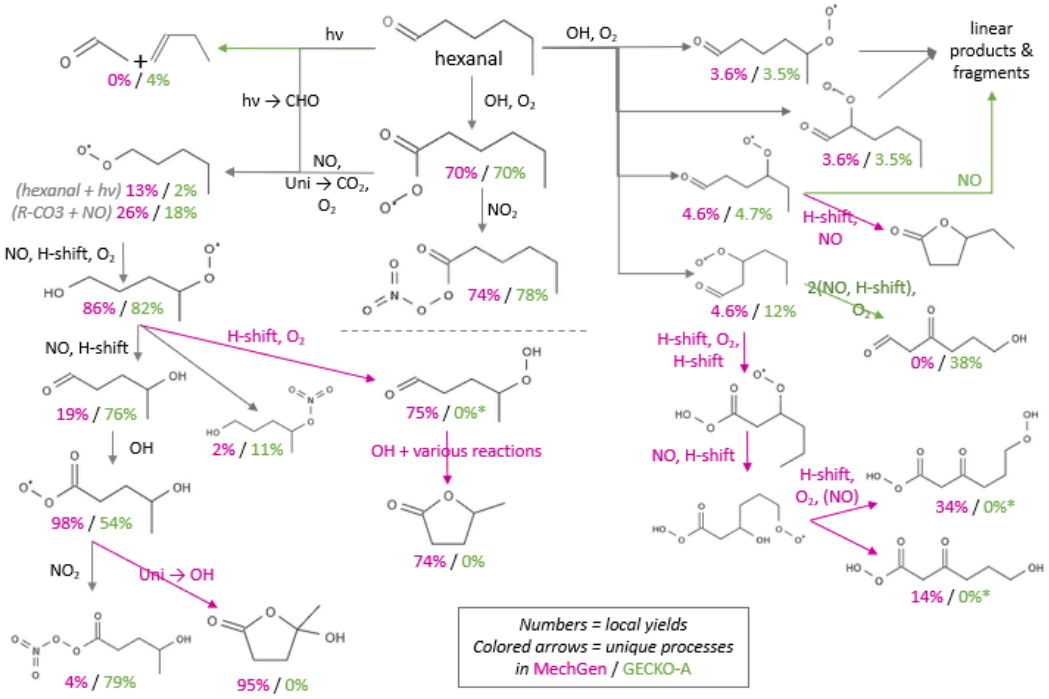
Hexanal loss under the specified conditions is dominated by reaction with OH, with a minor (6% in GECKO-A, 13% in MechGen) contribution from photolysis (reaction scheme 5). The strongest influence on NDPOHR, and likely also on fragmentation, is the production once more of unreactive (with relatively low kOH, $< 5 \times 10^{-12}$ cm^3 $molecule^{-1}$ s^{-1}) and therefore persistent lactones in MechGen but not in GECKO-A. The 1st- and 4th-most abundant MechGen products are two lactones originating from the 70% OH reaction channel (center and left of reaction scheme 5), while the 3rd-most abundant is produced from a 4.6% OH reaction channel (upper right of reaction scheme 5).

Another interesting feature is the occurrence in GECKO-A of Norrish type II photolysis of hexanal (upper left side of reaction scheme 5), a process not included in MechGen. Hexanal photolyzes to acetaldehyde (kOH = 1.5×10^{-11} cm^3 $molecule^{-1}$ s^{-1}) and but-1-ene, which oxidizes rapidly to propanal (kOH = 4.4×10^{-11} cm^3 $molecule^{-1}$ s^{-1}). This contributes to GECKO-A's two-times higher acetaldehyde and PAN levels than in MechGen.

We note for completeness that, for hexanal chemistry, the use in GECKO-A of a branching ratio threshold of 5% (the open-source default, Aumont et al., 2026) rather than the 0.5% used herein excludes three of the five initial OH+hexanal reaction channels, redistributing their mass flux into the major (70%, RCO_3) channel and one minor channel (the 12% channel on the right of Reaction Scheme 5). Similar first-generation minor-pathway exclusion would be found for OH reactions with the CH_3 groups in octane or with certain CH_2 groups in longer n-aldehydes, where the branching ratios are just below 5%. We performed a sensitivity test with n-octane using a 5% branching ratio threshold, and found that while the resulting mechanism was significantly reduced in size compared to the case with an 0.5% threshold, there was negligible effect on bulk simulation metrics under the conditions of the current

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

study.



Reaction scheme 5: Hexanal: major predicted photo-oxidation reactions. Numbers are time-integrated local yields in the MechGen (magenta) and GECKO-A (green) box model simulations. Stars indicate local yield by pathway shown: these species are also produced in GECKO-A by low-yield HO_2 pathways (not shown).

3.2.6. *trans*-2-Hexene

The early stages of *trans*-2-hexene (t-2-hexene) chemistry are characterized by large yields of butanal and its co-product acetaldehyde, via both OH- and ozone-initiated oxidation. The paired simulations yield some of the closest agreement in product mixture properties among our test molecules, with a mean agreement score of 0.8 (Table 3). Mean POHR, C:C_0 and O:C are almost identical, and Σ|Dfn|, *B(Pvap)*, and *B(H*)* are all within the "good-to-very good" agreement range. The two models have the same top 7 species (albeit in slightly different order) in terms of both abundance (table 9) and OHR rankings (table S9b), i.e. a range of small (C≤4) aldehydes and PAN-containing products. The close agreement between models, particularly in the early stages, is driven by the presence of the C=C double bond which provides a single dominant site for OH and ozone attack, with straightforward chemistry ensuing.

Table 9. Top 10 *t*-2-hexene products by simulation-integrated molecular abundance

Name or type	Canonical SMILES structure	Integrated abundance	
		GECKO-A	MechGen
PAN	CC(=O)OON(=O)=O	16%	20%
acetaldehyde	CC=O	14%	20%
formaldehyde	C=O	13%	6%
C4-PAN	CCCC(=O)OON(=O)=O	12%	8%
PPN	CCC(=O)OON(=O)=O	9%	8%

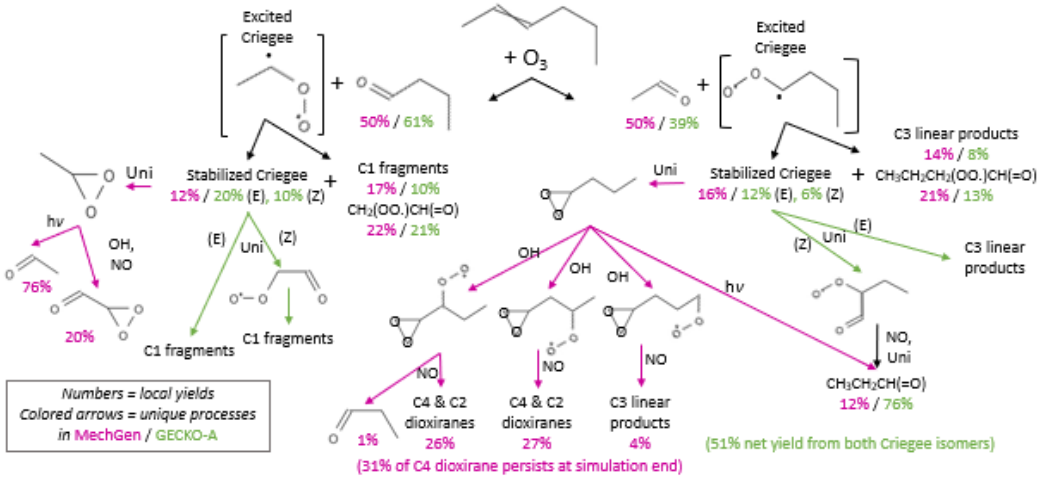
https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

propanal	CCC=O	6%	7%
butanal	CCCC=O	6%	6%
methanol	CO	3%	( )
a C4 stabilized Criegee/dioxirane	CCC[C]O[O] or CCC1OO1	-	3%
a C2 stabilized Criegee/dioxirane	C[C]O[O] or CC1OO1	-	3%
a C4 keto PAN	CC(=O)CC(=O)OON(=O)=O	2%	( )
propane	CCC	( )	2%
3-nitrooxy-hexan-2-one	CCCC(C(O)C)ON(=O)=O	1%	( )
Other products		19%	16%

( ) indicates species outside the top 10 in this simulation
- indicates species not produced in this mechanism

The minor (≤3% molecular abundance) products of t-2-hexene illustrate, however, some systematic inter-mechanism differences. The initial oxidation step produces two hydroxy-RO_2 intermediates, which each react with NO to yield a reactive C6 hydroxy-alkoxy radical, and small quantities of a C6 hydroxy nitrate. GECKO-A estimates ~30% higher branching ratios and 18% slower OH reaction rates for these hydroxy nitrates than does MechGen, resulting in their ~50% higher abundance in GECKO-A. (The more abundant isomer appears in Table 9). Also, in the second generation GECKO-A produces ethene as a minor product via Norrish type II photolysis of butanal, as already noted for hexanal (reaction scheme 5), while MechGen produces, via autoxidation, glyoxylic peracid which is ~30% more OH-reactive than ethene and is also less volatile and more soluble than butanal by 9 and 4 orders of magnitude, respectively.

Ozonolysis (reaction scheme 6), which represents ~40% of the net loss of t-2-hexene in these simulations, also shows inter-model differences. Ozonolysis has two pathways, each producing an aldehyde fragment and an excited ("hot") Criegee radical intermediate, which is considered too short-lived to be explicitly represented in either GECKO-A or MechGen as it immediately converts to a stabilized ("cold") Creigee intermediate and a range of decomposition products. The products in reaction scheme 6 reflect the simulation conditions of 0% RH, and we note that different pathways and products would be relevant at higher RH. For the current conditions, in GECKO-A, the cold Criegee radicals in turn either add oxygen to form a peroxy carbonyl (in the case of the Z-conformer), or form dioxiranes which rapidly rearrange and fragment (in the case of the E-conformer) (Newland *et al.*, 2022, and references therein). MechGen does not differentiate between conformers, and converts the cold Criegee radicals exclusively to the analogous dioxiranes, which the MechGen SAR predicts to be fairly long-lived. The linear portions of these dioxiranes react exclusively and relatively slowly with OH, so these species persist and appear among the most-abundant MechGen products in our simulation (*Table S9a*).



https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

Reaction Scheme 6: T-2-hexene: ozonolysis. Numbers are time-integrated local yields in the MechGen (magenta) and GECKO-A (green) box model simulations. CO and CO_2 are generated as coproducts of each channel.

3.2.7. Isoprene

Isoprene initial chemistry is considerably more challenging to fully represent than that of the other species considered here, owing to the presence of reversible reactions between hydroxyalkyl and hydroxy alkyl peroxy radicals in the first oxidation step (Wennberg *et al.*, 2018), pathways not included in either the GECKO-A or MechGen standard protocols. To overcome this challenge, GECKO-A includes and recommends an observationally-based isoprene chemistry option, designated here as "GECKO-MCM", that relies for several chemical generations upon handwritten chemistry from the MCM 3.3.1 model (Jenkin *et al*., 2015) and includes the reversible $R+O_2$ chemistry referred to above. This MCM-derived chemistry provides a 'benchmark' against which the two SAR-generated mechanisms can be compared, in addition to their comparison against each other. The top 10 most-abundant products from each of the three systems are provided in Table 10, and a subset of the chemistry is given for illustrative purposes in Reaction Scheme 7.

Table 10. Top 10 isoprene products by simulation-integrated molecular abundance

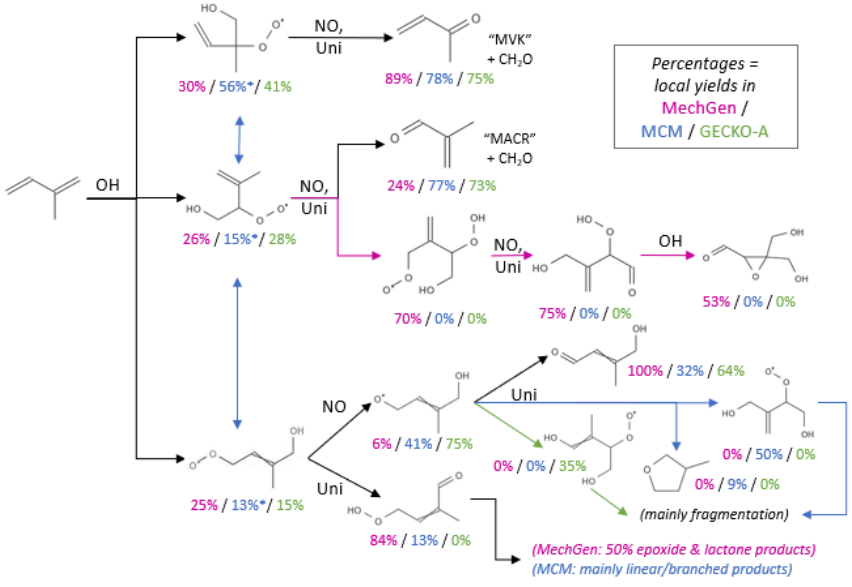
Name or type	Canonical SMILES structure	Integrated abundance		
		GECKO-SAR	GECKO-MCM	MechGen
formaldehyde	C=O	20%	26%	11%
PAN	CC(=O)OON(=O)=O	13%	14%	14%
hydroxyacetone	CC(=O)CO	11%	7%	3%
glycolaldehyde	OCC=O	7%	5%	5%
hydroxy-PAN	OCC(=O)OON(=O)=O	5%	4%	7%
methyl vinyl ketone (MVK)	CC(=O)C=C	5%	6%	4%
methyl glyoxal	O=CC(=O)C	3%	4%	2%
4-hydroxy-3-nitrooxy-butan-2-one	OCC(C(=O)C)ON(=O)=O	( )	7%	( )
glyoxal	O=CC=O	3%	( )	( )
di-hydroxy acetone	OCC(=O)CO	( )	( )	3%
3-methyl-3-hydroxy-oxolan-2,4-dione	O=C1OCC(=O)C1(C)O	-	-	3%
a C5 isoprene epoxydiol (IEPOX)	OCC1OC1(C)CO	-	2%	2%
methacrolein (MACR)	CC(=C)C=O	2%	( )	( )
3-hydroxy-4-nitrooxy-butan-2-one	O=N(=O)OOCC(C(=O)C)O	( )	2%	( )
methyl hydroperoxide	COO	2%	-	( )
Other products		30%	23%	45%

( ) indicates species outside the top 10 in this simulation
- indicates species not produced in this mechanism

In the SAR-based mechanisms (GECKO-A and MechGen), the branching ratios and radical products of the initial oxidation step are constrained following Paulot *et al.* (2009) and Carter *et al*. (2025), respectively. Each model follows its own SAR protocols for the subsequent steps. OH addition occurs at either of the double bonds, sometimes with double-bond migration (allylization) to form hydroxy-substituted RO_2. There are 4 major and 2 minor channels; chemistry associated with three of the major channels is shown in Reaction Scheme 7 along with branching ratios to three of the four main RO_2 intermediates. In GECKO-A these channels lead largely to the established major first-generation products methyl vinyl ketone and methacrolein ("MVK" and "MACR"), in addition to other fragmentation products. In the GECKO-MCM mechanism, the reversible nature of the initial steps particularly favors formation of the OCC(C(=C)C)O[O] radical for the conditions used here, resulting in enhanced net yields of MVK (top row of Reaction Scheme 7). Allylized radicals are formed in smaller yields in all mechanisms and result in an array of multifunctional branched five-carbon products (bottom rows of Reaction Scheme 7).

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

In MechGen and the MCM, each of the major first-generation products faces competition from autooxidation channels, with only a minor effect on product yields in the MCM but a major impact on MACR production in MechGen. Resultant abundance rankings for MACR are 9 in GECKO-SAR, 11 in GECKO-MCM, and 16 in MechGen (*table S10a*). Autoxidation also operates in several other MechGen chemical pathways, including the 25% pathway that proceeds via the [O]OCC=C(CO)C radical. Accordingly, while GECKO-SAR produces unsaturated hydroxy carbonyls (e.g. reaction scheme 7, 4th row), MechGen yields in addition several hydroperoxy-carbonyls (e.g. reaction scheme 7, bottom row) which give rise to substituted epoxide products, key reactive intermediates in the formation of isoprene-derived SOA (e.g. Wennberg et al, 2018), and lactones. Substituted epoxides are also produced in HO_2 reaction channels in the MCM and MechGen, however they appear in GECKO-SAR only as products of isoprene ozonolysis.



Reaction scheme 7: Isoprene: Selected pathways in the first generation of OH-initiated oxidation. Numbers indicate simulation-integrated yields from the previous species in MechGen (magenta), GECKO-MCM (blue), and GECKO-A (green). Double-headed arrows indicate rapid interconversion between RO_2 radicals in GECKO-MCM. Stars indicate simulation-integrated yields after interconversion, calculated relative to isoprene loss.

The isoprene product mixture bulk property metrics reflect these pronounced mechanistic differences between MechGen and either version of GECKO-A. GECKO-A favors fragmentation, yielding significantly higher quantities of CH_2O and other relatively reactive fragmentation products (Tables 10, S10a,b), whereas MechGen produces relatively unreactive epoxides and lactones. Consequently both GECKO-SAR and GECKO-MCM show significantly higher late-simulation POHR than MechGen (up to +60%) and poor-to-poorest C:C_0 agreement (table 3, figure S2). The ND(O:C) values suggest "good" atomic ratio agreement, however the O:C ratio is not specific enough to show the overall disparity in functionality ($\Sigma|Dfn|$ = 0.39, "moderate", and 0.51, "poor", for GECKO-MCM and GECKO-SAR, respectively). While GECKO-SAR preferentially generates aldehydes and PANs, MechGen emphasizes ketones, hydroperoxides, and peroxyacids (Figure 7, Figure S1). MechGen also generates up to 0.1 g:C each of ethers and esters, and 0.1 aliphatic rings per product carbon, including both epoxides and other persistent and unique cyclic products from radical cyclization, all of which are virtually absent in GECKO-SAR. GECKO-MCM produces marginally more epoxides and rings than GECKO-SAR, but far fewer than MechGen; conversely its yields of PAN and aldehydes are lower than in GECKO-A, but still greater than in MechGen. Also, GECKO-MCM generates lower quantities of ethers (0.013 g:C) than MechGen (0.07 g:C).

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

The extremely poor inter-model agreement in both volatility and solubility (figure 5) is a direct result of the divergent product mixtures. It is driven partly by the absence of autooxidation in either flavor of GECKO (which yields, in MechGen, products such as the hydroperoxy ketone CC(=CCOO)C=O, bottom line of reaction scheme 7) and partly by MechGen's production of persistent functionalized cyclic species such as the epoxide OCC1(CO)OC1C=O (reaction scheme 7, middle right). The *Pvap* values for these two products are relatively low, at 5.6×10^{-5} and ~ 3×10^{-7} atm, while their *H** values are relatively high, at 2.3×10^{5} M.atm^{-1} and 2.6×10^{9} M.atm^{-1}, respectively (compare to the *Pvap* and *H** distributions, figures 3 and 4). An analysis of product functionality distribution in *Pvap* and *H** space (Figure S5) shows that mono- and difunctional species such as MVK (line 6 of table 10, denoted "UK" in Figure S5) which are more prevalent in GECKO-A increase overall *Pvap* and lower *H**, while cyclic multifunctional species (especially epoxides and lactones, more prevalent in MechGen, denoted T*E and T*J in Figure S5) lower overall *Pvap* and raise *H**.

3.2.8. α-Pinene

Despite its molecular complexity, the GECKO-A and MechGen model results for α-pinene initially agree moderately well. OHR is driven by the dominance of the precursor and is essentially identical in the two systems. NDPOHR is within ± 12.5% for most of the first 2 hours (~8 precursor lifetimes), however agreement decreases with increasing oxidation time, giving a 6-hour mean of 30% ± 25%. The model systems show visually similar vapor pressure and solubility distributions for the first 5 precursor lifetimes: *B(Pvap)* is statistically insignificant, while *B(H*)* (0.44) is in the "moderate" range. Whole-simulation ND(O:C) and ND(C:C_0) are 0.17 (poor) and -0.5 (moderate), respectively, indicating more oxygenation and more fragmentation in GECKO-A. Whole-simulation functionality agreement is relatively good (Σ|Dfn| = 0.19). The most pronounced functional group differences between the two models are in MechGen's production of persistent γ-lactones, and in GECKO-A's preference for multifunctional nitrates and PANs. As seen with previous precursors, the γ-lactones originate from autooxidation and radical cyclization reactions that occur in MechGen but not GECKO-A. One example is O=C1CC(C(O1)(C)C)C(=O)C (Table 11, line 5).

Table 11. Top 10 α-pinene products by simulation-integrated molecular abundance

Name or type	Canonical SMILES structure	Integrated abundance	
		GECKO-A	MechGen
acetone	CC(=O)C	18%	17%
formaldehyde	C=O	11%	3%
PAN	CC(=O)OON(=O)=O	9%	18%
a C10 (cyclic) keto PAN	O=C(CC1CC(C1(C)C)C(=O)C)OON(=O)=O	4%	3%
a C8 acyl-substituted lactone	O=C1CC(C(O1)(C)C)C(=O)C	( )	3%
a C10 dioxirane	CC(=O)C1CC(C1(C)C)CC2OO2	-	3%
pinonaldehyde	O=CCC1CC(C1(C)C)C(=O)C	2%	3%
a C9 (cyclic) keto PAN	O=N(=O)OOC(=O)C1CC(C1(C)C)C(=O)C	2%	2%
a C3 di-PAN	O=C(CC(=O)OON(=O)=O)OON(=O)=O	2%	( )
oxolan-2,4-dione	O=C1OCC(=O)C1	-	2%
a C10 bicyclic hydroxy nitrate	O=N(=O)OC1CC2CC(C1(C)O)C2(C)C	1%	( )
a C10 bicyclic hydroxy ketone	O=C1CC2CC(C1(C)O)C2(C)C	1%	( )
a C5 keto di-PAN	O=C(CC(=O)OON(=O)=O)CC(=O)OON(=O)=O	1%	( )
a C9 (cyclic) keto aldehyde	O=CC1CC(C1(C)C)C(=O)C	( )	1%
Other products		49%	44%

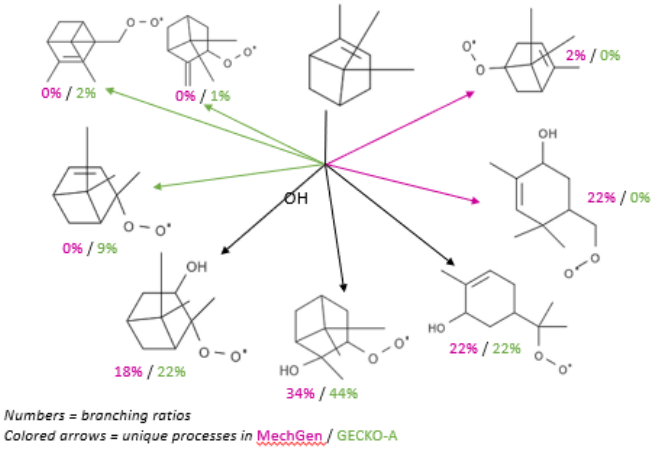
( ) indicates species outside the top 10 in this simulation
- indicates species not produced in this mechanism

Reaction with OH accounts for about ⅔ of the a-pinene loss in both systems, with ozone reaction accounting for the remainder. The majority of the OH reaction occurs via addition to either side of the double bond, forming two

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

hydroxy alkyl radicals. Ring-opening can occur for one of these radicals (see Reaction Scheme 8) - this occurs automatically via SARs in MechGen, with an overall extent of 43% of the OH products, while GECKO-A uses handwritten chemistry at this stage (following Peeters *et al*., 2001), with an overall ring-opening yield for of 22% of the OH products. Note that GECKO-A specifies only one ring-opened product, while MechGen forms two possible species, estimated to have roughly equal yields. GECKO-A also includes an abstraction pathway (with 9% branching ratio) that, following allylization, leads to an unsaturated peroxy radical (see reaction scheme 8).

Despite some of these differences in initial reaction channels, major products formed in the two systems show significant overlap – for example, five of the top ten products in terms of OHR in each mechanism are shared (*table S11b*). Examples include α-pinene itself, formaldehyde, pinonaldehyde, and a C9 product of pinonaldehyde. In terms of abundance, acetone and PAN are two of the three top molecules formed in both systems (*Table 11*). As highlighted for t-2-hexene, MechGen predicts that ozonolysis of *α*-pinene results in unique and highly-stable dioxirane species (via the stabilized Criegee route), while in GECKO-A the stabilized Criegees give rise to numerous substituted fragmentation products.



Reaction scheme 8: α-Pinene: RO_2 products in the first-generation of OH-initiated α-pinene oxidation. Numbers are OH-reaction-specific branching ratios in MechGen (magenta) and GECKO-A (green).

3.2.9. Camphene

Results for **camphene** show moderately close agreement between the two models in the first 5 precursor lifetimes (~2 hours), with statistically insignificant *B(Pvap)*, *B(H*)*, and ND(C:C_0) values, and "good" (<26%) NDPOHR (Figure 2). On the whole-simulation (6-hour) timescale, the models diverge: ND(C:C_0) remains statistically insignificant, product functionality shows only moderate simulation-mean divergence (Σ|Dfn| = 0.21), but disagreement in O:C and POHR increases.

The major products generated by MechGen and GECKO-A in the oxidation of camphene are somewhat similar, as evidenced by the similar lists of top 10 products both in terms of integrated abundance and integrated OH reactivity over the timescale of the box model runs (*tables 12 and S12a,b*). Major species generated by both mechanisms include, as expected, 'traditional' species such as acetone and formaldehyde, and first- and second-generation hydroxynitrates and hydroxy-hydroperoxides. The similarity stems from a number of factors that include 1) the existence of one major site for OH attack (at the external =CH2 group) that pushes the majority of the chemistry down a single path, 2) similar rules for alkoxy radical isomerizations in the two systems that drive much of the chemistry for this species, 3) slow ring-opening upon OH addition in either system (the rings are deemed too large by MechGen for this to be a major pathway) and 4) inhibition of autoxidation until after ring-breaking, due to the bicyclic structure of the camphene backbone. (The rigidity of the molecular structure leads to the logical assumption

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

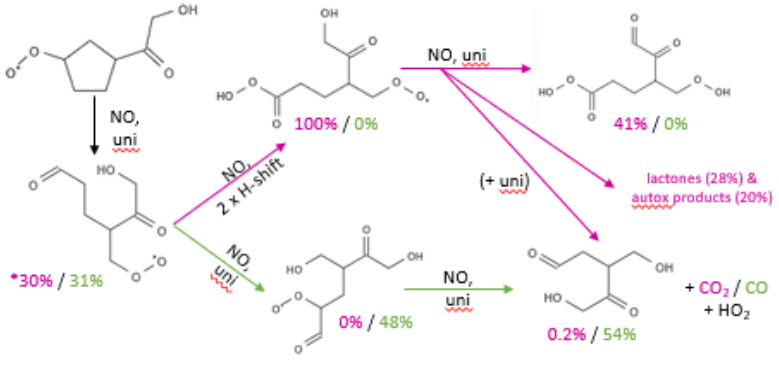
that three carbon atoms belonging to the same ring would not participate in autoxidation). In many ways then, the chemistry of camphene in the initial stages follows that of a relatively simple alkene.

Table 12. Top 10 camphene products by simulation-integrated molecular abundance

Name or type	Canonical SMILES structure	Integrated abundance	
		GECKO-A	MechGen
acetone	CC(=O)C	30%	38%
formaldehyde	C=O	8%	3%
hydroxyacetyl peroxynitrate	OCC(=O)OON(=O)=O	5%	5%
a C10 bicyclic hydroxy nitrate	OCC1(ON(=O)=O)C2CCC(C1(C)C)C2	3%	3%
oxolan-2,5-dione	O=C1CCC(=O)O1	-	3%
a C10 (5-cyclic) hydroxy keto nitrate	OCC(=O)C1CCC(C1)C(ON(=O)=O)C	2%	2%
a C10 bicyclic hydroxy hydroperoxide	OCC1(OO)C2CCC(C1(C)C)C2	2%	2%
a C9 bicyclic ketone	O=C1C2CCC(C1(C)C)C2	2%	2%
a C7 (5-cyclic) hydroxy keto nitrate	OCC(=O)C1CCC(C1)ON(=O)=O	2%	( )
a C5 hydroperoxy keto peracid	OOCCC(=O)CC(=O)OO	( )	2%
a C7 (5-cyclic) hydroxy di-ketone	OCC(=O)C1CCC(=O)C1	1%	3%
a C10 (5-cyclic) hydroxy hydroperoxy ketone	OCC(=O)C1CCC(C1)C(OO)(C)C	1%	( )
Other products		43%	37%

( ) indicates species outside the top 10 in this simulation
- indicates species not produced in this mechanism

After the first camphene ring breaks, both mechanisms predict attack by NO on the resulting single-ring substituted RO_2. The subsequent chemistry of the branched product shows the familiar differences in approach including competitive autoxidation and recyclization, resulting in diverging later-stage product mixture properties. A specific example is shown in Reaction Scheme 9.



Reaction Scheme 9: Camphene: Selected reactions of the major (30% net yield) C7 RO_2 product.
The asterisk indicates an implicit step in the MechGen scheme. Numbers indicate simulation-integrated local yields. The GECKO-A-specific RO_2 intermediate (lower line) is one of two equal-yield isomers.

3.2.10. 2-Methyl furan (2-mf).

Model output from GECKO-A and MechGen for 2-methyl furan ("2-mf") yields five of the same top 10 most-abundant species (table 13) including large amounts of acetone and some formaldehyde, followed by a range of diverse substituted cyclic and linear products. Even among the shared products, however, the relative proportions

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

vary considerably, potentially leading to substantial differences in product mixture properties. The lists of highest OHR-contributors have similarly poor overlap (table S13b), indicating property divergence among lower-abundance but more OH-reactive products. The discussion that follows refers to version 0 of GECKO-A (Aumont et al., 2026) as the "standard" version, and the version with k(OH + maleic anhydride) set to that used in MechGen as the "sensitivity" version, or GECKO-A*, with the results presented in the order "standard" followed by "sensitivity".

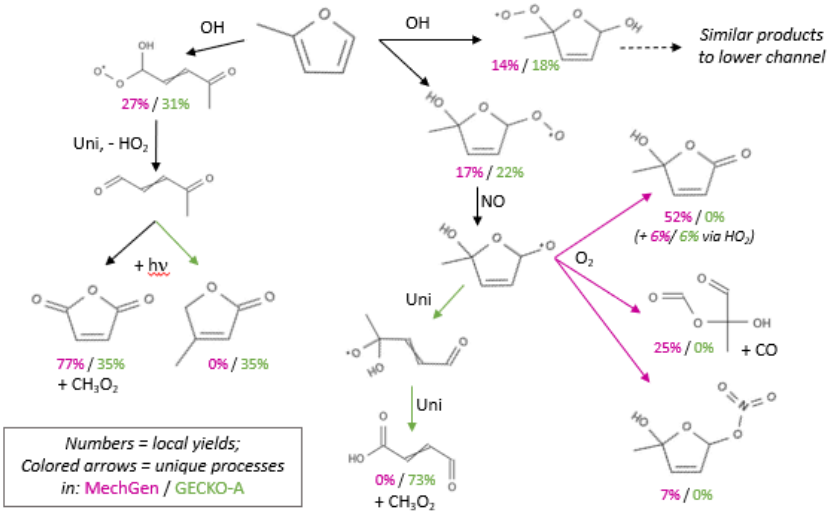
2-Methyl furan shows the largest inter-model difference in bulk product functionality of all our precursors, with $\Sigma|Dfn| = 0.82$ and 0.54 for the standard and sensitivity versions of GECKO-A, respectively. Accordingly, the physical property agreement score is among the largest, at 2.7 (2.3 for GECKO-A*, Table 3). NDPOHR, ND(O:C) and ND(C:C_0) average 0.88 ("poorest"), 0.23 ("poor") and -0.15 ("poor"), respectively, over the 6 hour simulation, i.e. the GECKO-A product mixture is much more OH-reactive, contains considerably more oxygen, and undergoes more fragmentation than does the MechGen mixture. The ratio values improve to 0.37 ("good") and 0.11 ("moderate") in the sensitivity case, however NDPOHR increases to 1.03 ("Poorest"). While *B(Pvap)* and *B(H*)* are considered "good" and "moderate" at 0.28 and 0.47, their values depart substantially from the expected anticorrelation, giving the largest δc in this study, 0.53 (Fig 5). For GECKO-A*, *B(Pvap)* = 0.24, *B(H*)* = 0.37, and $\delta c = 0.43$. This suggests inter-model chemistry differences beyond those of our other study precursors, which are not substantially resolved in the sensitivity study.

Table 13. Top 10 2-methyl furan products by simulation-integrated molecular abundance.

Name or type	Canonical SMILES structure	Integrated abundance		
		GECKO-A	GECKO-A*	MechGen
formaldehyde	C=O	17%	16%	5%
PAN	CC(=O)OON(=O)=O	9%	8%	5%
maleic anhydride	O=C1C=CC(=O)O1	3%	15%	23%
methyl carbonate	O=COC(=O)C	3%	3%	12%
acetic acid	CC(=O)O	6%	6%	2%
glyoxal	O=CC=O	6%	4%	( )
a C3 hydroxy aldehyde acylperoxynitrate	O=CC(C(=O)OON(=O)=O)O	3%	1%	( )
a C3 hydroxy di-acylperoxynitrate	OC(C(=O)OON(=O)=O)C(=O)OON(=O)=O	3%	( )	-
a C3 dicarbonate	O=C1OC(=O)C(=O)O1	-	-	3%
a C5 cyclic hydroxy epoxy ester	O=C1OC2(C(C1O)O2)C	-	-	3%
a C4 hydroxy keto acid	OC(C(=O)O)C(=O)C	2%	2%	-
a C4 unsaturated aldehydic acid	O=CC=CC(=O)O	2%	2%	-
methyl hydroperoxide	COO	( )	2%	( )
a C4 cyclic hydroxy carbonate	O=C1OC(=O)C(O1)(C)O	-	-	2%
a C4 ester acylperoxynitrate	CC(=O)OC(=O)OON(=O)=O	-	-	2%
a C4 cyclic diester nitrate	O=N(=O)OC1(C)OC(=O)OC1=O	-	-	1%
Other products		45%	40%	42%

GECKO-A* is the GECKO-A mechanism amended to use the k(OH + maleic anhydride) value from MechGen.
( ) indicates species outside the top 10 in this simulation.
- indicates species not produced in this mechanism.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026



766

767 **Reaction Scheme 10: 2-Methyl Furan: Selected reactions after OH-addition**. Percentages indicate
768 reaction-specific branching ratios in MechGen (magenta) and GECKO-A (green).

769 Reaction with OH is the major loss of 2-mf in both systems. The initial steps in the OH + 2-mf chemistry are
770 constrained following theoretical results from Yuan *et al*. (2017) such that the two generators start with essentially
771 the same chemical information. (In reaction scheme 10, these constraints apply only to the first and second arrows
772 on each side of the diagram). While both systems consider reaction of 2-mf with NO_3, the relative importance of this
773 chemistry is larger in GECKO-A than in MechGen. Only MechGen considers reactions with ozone.

774 The chemistry after OH addition to 2-mf follows each generator's specific protocols, and is conceptually similar in
775 both mechanisms. OH addition occurs exclusively at the 2- and 5- positions, leading down two general pathways.
776 The first pathway (left side of reaction scheme 10) is a unique ring-opening process giving an unsaturated
777 dicarbonyl species, similar to species formed in aromatic oxidation chemistry (albeit by different routes). The
778 second general pathway (right side of reaction scheme 10) forms ring-retaining allylized hydroxy-peroxy radicals
779 that lead to either substituted ring-retaining products or unsaturated dicarbonyls as first-generation end products.
780 Re-cyclization of unsaturated dicarbonyls via photolysis is also a common feature of both mechanisms. Despite
781 these general commonalities, however, there are many quantitative differences in branching ratios and hence yields
782 of even first- and second generation products. One example is the chemistry of the ring-retaining peroxy radical
783 species generated with a 17-22% yield from the initial 2-mf + OH reaction, followed in both models by conversion
784 via reaction with NO to the corresponding alkoxy radical. GECKO-A decomposes the alkoxy via ring-opening
785 leading exclusively to a multifunctional acid, whereas MechGen reacts it with O_2 forming a ring-opened dicarbonyl
786 ether with 25% yield, and several ring-retaining species including a hydroxy lactone with 52% yield (Reaction
787 Scheme 10). That quantitative differences exist in the two systems in the treatment of 2-mf derived alkoxy radicals
788 in particular is not surprising given their complexity (e.g., oxy radical functionality located β to ethers and C=C
789 double bonds), and the paucity of training data available to constrain SARs.

790 Overall OH reactivity is controlled by 2-mf itself with a small difference in the k(OH + 2-mf) between the two
791 mechanisms, 7.3×10^{-11} cm^3 molec^{-1} s^{-1} in GECKO-A vs 7.0×10^{-11} cm^3 molec^{-1} s^{-1} in MechGen. However, NDPOHR
792 for 2-mf is the largest among our precursors, reaching almost 120% later in the simulation. As noted by Aumont et
793 al (2026), the SARs used in GECKO-A yield OH rate coefficients for maleic anhydride, O=C1C=CC(=O)O1, of 1.7
794 $\times$ 10^{-11} cm^3 molec^{-1} s^{-1}, significantly faster than the measured value of 3.9×10^{-13} cm^3 molec^{-1} s^{-1} adopted in
795 MechGen. For its hydroxy analog, OC1C=CC(=O)O1, GECKO-A and MechGen use SARs which yield OH rate
796 coefficients differing by a factor of two: 8.7×10^{-11} and 4.1×10^{-11} cm^3 molec^{-1} s^{-1}, respectively. We performed a
797 sensitivity test by reducing the GECKO-A value for k(OH+maleic anhydride) to that used by MechGen. The

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

seemingly counterintuitive resulting increase in NDPOHR (Sect 3.1.1) arises mainly from increased maleic anhydride longevity (*tables S13a,b*). The SARs for maleic anhydride and similar compounds have been marked for future revision in GECKO-A (Aumont et al., 2026).

2-Mf is also notable in that the GECKO-A product mixture is somewhat more soluble than that in MechGen, with the most positive *B(H*)* value in this study (+0.47, classed as "moderate" agreement in comparison to isoprene's strongly negative value). This is at least partly due to the formation of acid groups, which confer relatively high solubility (see Figure S5) and are produced in GECKO-A by multi-stage radical processes following ring-opening. Acid groups approach a relative abundance of 0.1 groups per product carbon in GECKO-A, but only 0.02 groups per product carbon in MechGen, Figures 7, S1. Vapor pressures of the product mixtures are more visually similar, but distribution means are, again, moderately different (*B(Pvap)* +0.28) owing partly to the presence in MechGen of low-volatility (~10^{-10} atm) epoxides and lactones (see Figure S5).

Maleic anhydride is ~3-fold more prevalent in MechGen than in GECKO-A v0 at five 2-mf lifetimes, and is the single greatest contributor to the inter-model differences at 10^{-6} atm vapor pressure and 10^{2} M.atm^{-1} solubility (Figures 3, 4, and species "TUKKE" in figure S5). Its concentration in the GECKO-A sensitivity case increases to ~ 60% of the MechGen value, reducing but not eliminating the differences in these specific distribution bins. The 10^{-6} atm *Pvap* increase comes at the expense of the vapor pressure bins on either side, with reductions in concentrations of O=CC(C(=O)OON(=O)=O)O (species "PDO" in Figure S5) and of a variety of lower-abundance lower-volatility products, explaining why *B(Pvap)* hardly changes in the sensitivity case. *B(H*)*, while still positive in the sensitivity case, shows slightly larger reductions than *B(Pvap)* since the ~25% increase in the (relatively insoluble) maleic anhydride results mainly in comparable reductions in higher-solubility species. 2-Mf's large δc value (0.53), indicating non-conformance with the usual *B(pvap)/B(H*)* anticorrelation, is only partially reduced (to 0.43) in the sensitivity study. This, along with the only partial reduction in overall agreement score (from 2.7 to 2.3), illustrates numerically that systemic inter-model differences in the 2-mf mechanisms exist in addition to the maleic anhydride OH rate constant, such as those discussed above.

3. SUMMARY

Herein, we have compared the results of box model simulations for 10 VOCs under a standard set of conditions using mechanisms generated by the two most common nearly explicit mechanism generator systems for atmospheric VOC oxidative chemistry; GECKO-A and MechGen. The two systems are conceptually similar (essentially using SARs for each class of reaction to extend available data), but differences clearly exist, such as in the range of processes included, the level of detail included for certain processes, and the specific SARs being used. While it is then not surprising that differences are seen in the yields and identities of the major products resulting from mechanisms generated by the two systems, different outcomes in terms of bulk properties of the product distributions (e.g., functional groups, OH reactivity, Henry's Law coefficient distributions, and vapor pressure distributions) also appear in some cases.

The SARs used by the two systems for OH+VOC reactions are quite similar, and thus the initial branching to various peroxy radicals shows only small differences. GECKO-A includes a more detailed representation of photolysis processes: one noticeable manifestation of this is in regard to Norrish type-2 photolysis, where very reactive alkenes can be generated in GECKO-A from aldehyde photolysis. GECKO-A is lacking in the inclusion of some alkoxy radical isomerization processes (those not involving six-membered rings) and does not yet include any representation of autoxidation. This leads to differences in many outcomes, for example in the ratio of hydroperoxide to hydroxide groups for larger species, or in the formation (or lack thereof) of isoprene-derived epoxides. One process that occurs commonly in MechGen that is not considered in GECKO-A is the cyclization of specific .RC(O)OX radicals resulting in formation of persistent lactone / carbonate type species. While the theoretical results of Curran et al. (1998), used in MechGen, support competitive rates for these processes, other (more recent) theoretical studies (Miyoshi, 2011; Ali, 2021) suggest that they are considerably slower. Further consideration and study of these reaction pathways is clearly required.

These differences (and other more minor ones) in the two generators lead broadly to higher POHR, higher vapor pressures and lower solubility of products formed from GECKO-A-derived mechanisms, as well as a larger propensity for formation of PANs and alcohols in GECKO-A versus ketones and hydroperoxides in MechGen.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

The analyses conducted here show that, despite their complexity and ongoing efforts to include "all known chemistry", both models have limitations and work to improve them should (and will) continue. Aumont et al (2026) have summarized known limitations of GECKO-A, and listed ongoing or planned improvements for future versions. In particular, work is in progress to incorporate the ever-expanding knowledge of autoxidation pathways and rates into GECKO-A. These recent advances can also be used to improve the already-existing MechGen capability. 1,6-H shifts in alkoxy radical isomerizations, already included in MechGen and shown here to be relevant in long-chain alkane chemistry, are also being added to GECKO-A. While an update is planned to the GECKO-A photolysis protocol to incorporate recent advances, overall knowledge of photolysis rates (absorption spectra, and particularly quantum yields) is limited for complex organics, and further lab work and SAR development in this area is encouraged.

Finally, the ultimate goal of quantitatively matching output of the detailed mechanism generators with the burgeoning ability of instrumentation to measure the vast array of partially oxygenated products formed in the atmosphere and/or in chamber experiments is an ongoing and exciting challenge.

Code availability

The GECKO-A v1.0 source code is permanently archived on Zenodo at https://doi.org/10.5281/zenodo.15309904 (Aumont et al., 2025), and is publicly available on GitLab (https://gitlab.in2p3.fr/ipsl/lisa/geckoa/public). The GECKO-A box model (NSF NCAR version) used in this work is archived on Zenodo at https://doi.org/10.5281/zenodo.21815388 (Lee-Taylor et al., 2026a) and is publicly available on GitHub at (*publication forthcoming: anticipated before this manuscript is published*), with instructional wiki pages. The source code for a version of MechGen later than that used in this study, but which employs the same SARs and chemical assumptions, is archived on Zenodo at https://doi.org/10.5281/zenodo.16622705 (Carter et al., 2025c) and is publicly available on GitHub (https://github.com/SAPRC/MechGen).

Data availability

All the mechanisms presented here and the box model output are permanently archived on Zenodo at https://doi.org/10.5281/zenodo.21813881 (Lee-Taylor et al., 2026b). The MechGen mechanisms are supplied in both their original form and as reformatted for use with the GECKO-A box model used at NSF NCAR. All mechanisms are supplied in ascii format and are also incorporated into the NetCDF-format box model output files.

Author Contributions

J.L-T., J.J.O., W.P.L.C., and K.C.B. designed the study, performed the simulations, and wrote the manuscript. B.A., R.V., M.C., and J.L-T developed the GECKO-A v1.0 code, and, with J.J.O., developed the GECKO-A protocols. W.P.L.C. developed the MechGen code and protocols. J.L.-T. developed the MechGen-to-GECKO translation protocol. J.L.-T. and Z.W. performed analysis and produced the Figures. All authors participated in the review and editing of the manuscript.

Competing Interests

At least one of the (co-)authors is a member of the editorial board of Atmospheric Chemistry and Physics.

Acknowledgements

The authors are grateful to Sasha Madronich for his pioneering contributions to the GECKO-A project. We also thank the many collaborators who have contributed to the development of GECKO-A over the past two decades, including Camille Mouchel-Vallon for his contributions to the GECKO-A box model used at NSF NCAR. J.L-T. gratefully acknowledges Hamish McMillen for assistance with geometric equations.

We would like to acknowledge computing support from the Casper system (https://ncar.pub/casper) provided by the NSF National Center for Atmospheric Research (NCAR), which is a major facility sponsored by the U.S. National Science Foundation under Cooperative Agreement No. 1852977.

This work was supported by a grant from the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) program, Assistance Agreement No. 84000701, awarded by the U.S. EPA. It has not been formally reviewed by EPA. EPA does not endorse any products or commercial services mentioned in this publication.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

References

Ali, M.A. (2021). Ab initio rate coefficients for reactions of 2,5-dimethylhexyl isomers with O2: temperature- and pressure-dependent branching ratios, *Phys. Chem. Chem. Phys.*, 2021,23, 6225-6240, https://doi.org/10.1039/D0CP06562E.

Atkinson, R. (2007). Rate constants for the atmospheric reactions of alkoxy radicals: An updated estimation method, *Atmos. Environ.*, 41(38), 8468-8485, https://doi.org/10.1016/j.atmosenv.2007.07.002.

Aumont, B., Szopa, S., & Madronich, S. (2005). Modelling the evolution of organic carbon during its gas-phase tropospheric oxidation: development of an explicit model based on a self-generating approach, *Atmos. Chem. Phys.*, 5, 2497-2517, https://doi.org/10.5194/acp-5-2497-2005.

Aumont, B., Valorso, R., Rickard, A., Jenkin, M., Lee-Taylor, J., Orlando, J.J., Madronich, S., Vereecken, L., & Camredon, M. (2026). GECKO-A v1.0: Exploring VOC Oxidation Trajectories Through Comparison with the Master Chemical Mechanism, in review, *Geoscientific Model Development (GMD)*, https://doi.org/10.5194/egusphere-2026-2907.

Camredon, M., Aumont, B., Lee-Taylor, J., & Madronich, S. (2007). The SOA/VOC/NOx system: an explicit model of secondary organic aerosol formation, *Atmos. Chem. Phys.*, 7, 5599-5610, https://doi.org/10.5194/acp-7-5599-2007.

Carter, W. P. L. (2000): "Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment," Report to the California Air Resources Board, Contracts 92-329 and 95-308, May 8. Available at https://www.cert.ucr.edu/~carter/absts.htm#saprc99.

Carter, W. P. L. (2021). Estimation of Rate Constants for Reactions of Organic Compounds under Atmospheric Conditions. *Atmosphere*, *12*(10), 1250. https://doi.org/10.3390/atmos12101250.

Carter, W. P. L., Jiang, J., Orlando, J. J., & Barsanti, K. C. (2025a). Derivation of atmospheric reaction mechanisms for volatile organic compounds by the SAPRC mechanism generation system (MechGen), *Atmos. Chem. Phys*., 25, 199-242, https://doi.org/10.5194/acp-25-199-2025.

Carter, W. P. L., Jiang, J., Wang, Z., and Barsanti, K. C. (2025b). The SAPRC Atmospheric Chemical Mechanism Generation System (MechGen), Geosci. Model Dev., 18, 8461-8483, https://doi.org/10.5194/gmd-18-8461-2025.

Carter, W. P. L (2025c), SAPRC Mechanism Generation System for the Atmospheric Reactions of Volatile Organic Compounds in the Presence of NO_x, https://intra.engr.ucr.edu/~carter/MechGen/ (last access: 25 July 2025).

Compernolle, S., Ceulemans, K., & Müller, J. F. (2010). Vapor pressure estimation methods applied to secondary organic aerosol constituents from α-pinene oxidation: an intercomparison study. *Atmos. Chem. Phys.*, 10(13), 6271-6282, https://doi.org/10.5194/acp-10-6271-2010.

HJ Curran, H.J. , Gaffuri, P., Pitz, W/J., & Westbrook, C.K. (1998). A comprehensive modeling study of n-heptane oxidation. *Combust. Flame*, 114 (1-2), 149-177, https://doi.org/10.1016/S0010-2180(97)00282-4.

Hodzic, A., Aumont, B., Knote, C., Lee-Taylor, J., Madronich, S., & Tyndall, G. (2014), Volatility dependence of Henry's law constants of condensable organics: Application to estimate depositional loss of secondary organic aerosols, *Geophys. Res. Lett.*, 41(13), 4795-4804, https://doi.org/10.1002/2014GL060649.

Isaacman-VanWertz, G., & Aumont, B. (2021), Impact of organic molecular structure on the estimation of atmospherically relevant physicochemical parameters, *Atmos. Chem. Phys.*, 21(8), 6541–6563, https://doi.org/10.5194/acp-21-6541-2021.

Jenkin, M. E., Young, J. C., & Rickard, A. R. (2015). The MCM v3.3.1 degradation scheme for isoprene, *Atmos. Chem. Phys.*, 15, 11433-11459, https://doi.org/10.5194/acp-15-11433-2015.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

Jenkin, M. E., Valorso, R., Aumont, B., Rickard, A. R., & Wallington, T. J. (2018a). Estimation of rate coefficients and branching ratios for gas-phase reactions of OH with aliphatic organic compounds for use in automated mechanism construction, *Atmos. Chem. Phys.*, 18, 9297-9328, https://doi.org/10.5194/acp-18-9297-2018.

Jenkin, M. E., Valorso, R., Aumont, B., Rickard, A. R., & Wallington, T. J. (2018b). Estimation of rate coefficients and branching ratios for gas-phase reactions of OH with aromatic organic compounds for use in automated mechanism construction, *Atmos. Chem. Phys.*, 18, 9329-9349, https://doi.org/10.5194/acp-18-9329-2018.

Jenkin, M. E., Valorso, R., Aumont, B., & Rickard, A. R. (2019). Estimation of rate coefficients and branching ratios for reactions of organic peroxy radicals for use in automated mechanism construction, *Atmos. Chem. Phys.*, 19, 7691–7717, https://doi.org/10.5194/acp-19-7691-2019.

Jenkin, M. E., Valorso, R., Aumont, B., Newland, M. J., & Rickard, A. R. (2020). Estimation of rate coefficients for the reactions of O3 with unsaturated organic compounds for use in automated mechanism construction, *Atmos. Chem. Phys.*, 20, 12921-12937, https://doi.org/10.5194/acp-20-12921-2020.

Kerdouci, J., Picquet-Varrault, B., Durand-Jolibois, R., Gaimoz, C., & Doussin, J.-F. (2012). An Experimental Study of the Gas-Phase Reactions of NO3 Radicals with a Series of Unsaturated Aldehydes: trans-2-Hexenal, trans-2-Heptenal, and trans-2-Octenal, *J. Phys. Chem. A,* 116, 10135-10142, https://doi.org/10.1021/jp3071234.

La, Y. S., Camredon, M., Ziemann, P. J., Valorso, R., Matsunaga, A., Lannuque, V., Lee-Taylor, J., Hodzic, A., Madronich, S., & Aumont, B. (2016). Impact of chamber wall loss of gaseous organic compounds on secondary organic aerosol formation: explicit modeling of SOA formation from alkane and alkene oxidation, *Atmos. Chem. Phys.*, 16, 1417-1431, https://doi.org/10.5194/acp-16-1417-2016.

Lannuque, V., Camredon, M., Couvidat, F., Hodzic, A., Valorso, R., Madronich, S., Bessagnet, B., and Aumont, B. (2018), Exploration of the influence of environmental conditions on secondary organic aerosol formation and organic species properties using explicit simulations: development of the VBS-GECKO parameterization, *Atmos. Chem. Phys.*, 18(18), 13411–13428, https://doi.org/10.5194/acp-18-13411-2018.

Lee-Taylor, J., Madronich, S., Aumont, B., Baker, A., Camredon, M., Hodzic, A., Tyndall, G. S., Apel, E., & Zaveri, R. A. (2011). Explicit modeling of organic chemistry and secondary organic aerosol partitioning for Mexico City and its outflow plume, Atmos. Chem. Phys., 11, 13219–13241, https://doi.org/10.5194/acp-11-13219-2011.

Lee-Taylor, J., Hodzic, A., Madronich, S., Aumont, B., Camredon, M., & Valorso, R. (2015). Multiday production of condensing organic aerosol mass in urban and forest outflow. *Atmos. Chem Phys.*, *15*(2), 595-615, https://doi.org/10.5194/acp-15-595-2015.

Lee-Taylor, J.M., Aumont, B., Carter, W.P.L., Orlando, J.J., Wang, Z., Camredon, M., Valorso, R., & Barsanti, K.C. (2026a), Comparison of the GECKO-A and MechGen Atmospheric Chemical Mechanism Generators: Code Supplement (Version v0.0) [Dataset], https://doi.org/10.5281/zenodo.21815388.

Lee-Taylor, J.M., Carter, W.P.L., Aumont, B., Orlando, J.J., Wang, Z., Camredon, M., Valorso, R., & Barsanti, K.C. (2026b), Comparison of the GECKO-A and MechGen Atmospheric Chemical Mechanism Generators: Data Supplement (Version v0.0) [Dataset], https://doi.org/10.5281/zenodo.21813881.

Madronich, S., & Flocke, S. (1997). Theoretical Estimation of Biologically Effective UV Radiation at the Earth's Surface. In: Zerefos, C.S., Bais, A.F. (eds) Solar Ultraviolet Radiation. NATO ASI Series, vol 52. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-03375-3_3.

McVay, R. C., Zhang, X., Aumont, B., Valorso, R., Camredon, M., La, Y. S., Wennberg, P. O., and Seinfeld, J. H. (2016). SOA 1270 formation from the photooxidation of α-pinene: systematic exploration of the simulation of chamber data, Atmospheric Chemistry and Physics, 16, 2785–2802, https://doi.org/10.5194/acp-16-2785-2016.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

Miyoshi, A. (2011). Systematic Computational Study on the Unimolecular Reactions of Alkylperoxy (RO2), Hydroperoxyalkyl (QOOH), and Hydroperoxyalkylperoxy (O2QOOH) Radicals, *J. Phys. Chem. A*, 115, 15, 3301–3325, https://doi.org/10.1021/jp112152n.

Mouchel-Vallon, C., Lee-Taylor, J., Hodzic, A., Artaxo, P., Aumont, B., Camredon, M., Gurarie, D., Jimenez, J.-L., Lenschow, D. H., Martin, S. T., Nascimento, J., Orlando, J. J., Palm, B. B., Shilling, J. E., Shrivastava, M., & Madronich, S. (2021). Exploration of oxidative chemistry and secondary organic aerosol formation in the Amazon during the wet season: explicit modeling of 1285 the Manaus urban plume with GECKO-A, Atmos. Chem. Phys., 20, 5995–6014, https://doi.org/10.5194/acp-20-5995-2020.

Myrdal, P. B., & Yalkowsky, S. H. (1997). Estimating pure component vapor pressures of complex organic molecules. *Ind. & Eng. Chem. Res.*, 36(6), 2494-2499, https://doi.org/10.1021/ie9502421.

Nannoolal, Y., Rarey. J., Ramjugernath, D., & Cordes, W. (2004). Estimation of pure component properties: Part 1. Estimation of the normal boiling point of non-electrolyte organic compounds via group contributions and group interactions, *Fluid Phase Equilibria*, 226,45-63, https://doi.org/10.1016/j.fluid.2004.09.001.

Nannoolal, Y., Rarey, J., & Ramjugernath, D. (2008). Estimation of pure component properties: Part 3. Estimation of the vapor pressure of non-electrolyte organic compounds via group contributions and group interactions. *Fluid Phase Equilibria*, 269(1-2), 117-133, https://doi.org/0.1016/j.fluid.2008.04.020.

Newland, M. J., Mouchel-Vallon, C., Valorso, R., Aumont, B., Vereecken, L., Jenkin, M. E., & Rickard, A. R. (2022). Estimation of mechanistic parameters in the gas-phase reactions of ozone with alkenes for use in automated mechanism construction, *Atmos. Chem. Phys.*, 22, 6167–6195, https://doi.org/0.5194/acp-22-6167-2022.

O'Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., & Hutchison, G.R. (2011), Open Babel: An open chemical toolbox. *J. Cheminf. 3*, 33, https://doi.org/10.1186/1758-2946-3-33.

Pankow, J. F., & Asher, W. E. (2008). SIMPOL. 1: a simple group contribution method for predicting vapor pressures and enthalpies of vaporization of multifunctional organic compounds. *Atmos. Chem. Phys.*, 8(10), 2773-2796, https://doi.org/10.5194/acp-8-2773-2008.

Paulot, F., Crounse, J. D., Kjaergaard, H. G., Kroll, J. H., Seinfeld, J. H., & Wennberg, P. O. (2009). Isoprene photooxidation: new insights into the production of acids and organic nitrates. *Atmos. Chem. Phys.*, 9(4), 1479-1501, https://doi.org/10.5194/acp-9-1479-2009.

Peeters, J., Vereecken, L., and Fantechi, G. (2001). The detailed mechanism of the OH-initiated atmospheric oxidation of alpha-pinene: a theoretical study, Phys. Chem. Chem. Phys., 3, 5489–5504, https://doi.org/10.1039/b106555f.

Raventos-Duran, T., Camredon, M., Valorso, R., Mouchel-Vallon, C., & Aumont, B. (2010). Structure-activity relationships to estimate the effective Henry's law constants of organics of atmospheric interest, *Atmos. Chem. Phys.*, 10, 7643–7654, https://doi.org/10.5194/acp-10-7643-2010.

Valorso, R. , Aumont, B. , Camredon, M. , Raventos-Duran, T. , Mouchel-Vallon, C. , Ng, N. L., Seinfeld, J. H., Lee-Taylor, J. & Madronich, S. (2011). Explicit modelling of SOA formation from a-pinene photooxidation: sensitivity to vapour pressure estimation. *Atmos. Chem. Phys.*, 11, 6895-6910, https://doi.org/10.5194/acp-11-6895-2011.

Vereecken, L., & Peeters, J. (2009). Decomposition of substituted alkoxy radicals—part I: a generalized structure–activity relationship for reaction barrier heights. *Phys. Chem. Chem. Phys.*, 11(40), 9062-9074, https://doi.org/10.1039/B909712K.

Vereecken, L., & Peeters, J. (2010). A structure–activity relationship for the rate coefficient of H-migration in substituted alkoxy radicals, *Phys. Chem. Chem. Phys.*, 12, 12608-12620, https://doi.org/10.1039/C0CP00387E.

https://doi.org/10.5194/egusphere-2026-4908
Preprint. Discussion started: 21 August 2026

1017 Vereecken, L., Novelli, A., & Taraborrelli, D. (2017). Unimmolecular decay strongly limits the atmospheric impact
1018 of Criegee intermediates, *Phys. Chem. Chem. Phys.*, 47, https://doi.org/10.1039/C7CP05541B.

1019 Verwer, J. G. (1994). Gauss–Seidel Iteration for Stiff ODES from Chemical Kinetics, SIAM J. Sci. Comput., 15,
1020 1243–1250, https://doi.org/10.1137/0915076.

1021 Verwer, J. G., Blom, J. G., van Loon, M., & Spee, E. J. (1996). A comparison of stiff ODE solvers for atmospheric
1022 chemistry problems, Atmos. Environ., 30, 49–58, https://doi.org/10.1016/1352-2310(95)00283-5.

1023 Wallington, T., Ammann, K., Cox, R.A., Crowley, J.N., Herrmann, H., Jenkin, M.E., McHeill, V.F., Mellouki, A.W.
1024 & Troe, J., Evaluated Kinetic Data for Atmospheric Chemistry (2021), IUPAC,
1025 https://iupac.org/project/2017-024-1-100, accessed 2019.

1026 Wennberg, P.O., Bates, K. H., Crounse, J. D., Dodson, L. G., McVay, R. C., Mertens, L. A., Nguyen, T. B., Praske,
1027 E., Schwantes, R. H., Smarte, M. D., St. Clair, J.M., Teng, A. P., Zhang, X., & Seinfeld, J. H. (2018). Gas-phase
1028 reactions of isoprene and its major oxidation products, *Chem. Rev.*, 118, 3337-3390,
1029 https://doi.org/10.1021/acs.chemrev.7b00439.

1030 Yuan, Y., Zhao, X., Wang, S., & Wang, L. (2017). Atmospheric oxidation of furan and methyl-substituted furans
1031 initiated by hydroxyl radicals, *J. Phys. Chem. A*, 121, 9306–9319, https://doi.org/10.1021/acs.jpca.7b09741.