

RESEARCH ARTICLE

Automatic mechanism generation for the combustion of advanced biofuels: A case study for diethyl ether

Christian A. Michelbach  | Alison S. Tomlin

School of Chemical and Process Engineering, University of Leeds, Leeds, UK

Correspondence

Christian A. Michelbach, School of
Chemical and Process Engineering,
University of Leeds, Leeds, UK.
Email: c.michelbach@leeds.ac.uk

Funding information

Engineering and Physical Sciences
Research Council, Grant/Award Number:
EP/T033088/1

Abstract

Advanced biofuels have the potential to supplant significant fractions of conventional liquid fossil fuels. However, the range of potential compounds could be wide depending on selected feedstocks and production processes. Not enough is known about the engine relevant behavior of many of these fuels, particularly when used within complex blends. Simulation tools may help to explore the combustion behavior of such blends but rely on robust chemical mechanisms providing accurate predictions of performance targets over large regions of thermochemical space. Tools such as automatic mechanism generation (AMG) may facilitate the generation of suitable mechanisms. Such tools have been commonly applied for the generation of mechanisms describing the oxidation of non-oxygenated, non-aromatic hydrocarbons, but the emergence of biofuels adds new challenges due to the presence of functional groups containing oxygen. This study investigates the capabilities of the AMG tool Reaction Mechanism Generator for such a task, using diethyl ether (DEE) as a case study. A methodology for the generation of advanced biofuel mechanisms is proposed and the resultant mechanism is evaluated against literature sourced experimental measurements for ignition delay times, jet-stirred reactor species concentrations, and flame speeds, over conditions covering $\phi = 0.5\text{--}2.0$, $P = 1\text{--}100$ bar, and $T = 298\text{--}1850$ K. The results suggest that AMG tools are capable of rapidly producing accurate models for advanced biofuel components, although considerable upfront input was required. High-quality fuel specific reaction rates and thermochemistry for oxygenated species were required, as well as a seed mechanism, a thermochemistry library, and an expansion of the reaction family database to include training data for oxygenated compounds. The final DEE mechanism contains 146 species and 4392 reactions and in general, provides more accurate or comparable predictions when compared to literature sourced mechanisms across

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *International Journal of Chemical Kinetics* published by Wiley Periodicals LLC.

the investigated target data. The generation of combustion mechanisms for other potential advanced biofuel components could easily capitalize on these database updates reducing the need for future user interventions.

KEYWORDS

advanced biofuel, automatic mechanism generation, combustion, diethyl ether, oxygenated

1 | INTRODUCTION

Recent IPCC reports clearly state the urgent need to drastically reduce greenhouse gas (GHG) emissions and limit global warming to 1.5°C.^{1,2} However, global energy demand is predicted to grow significantly, due largely to a rise in global population and the increasing demands of developing regions for transportation and goods haulage.³ These sectors are heavily reliant on crude oil, raising environmental concerns which must be addressed rapidly. Alongside other factors (such as energy security and resource scarcity), these environmental concerns provide considerable motivation for the development of alternative (non-fossil) liquid fuels. Various such fuels have emerged with the potential to replace significant quantities of fossil-derived transportation fuels, including several advanced biofuels.^{4–9} Like all biofuels, advanced biofuels are produced from biomass feedstocks. However, unlike unsustainable first-generation biofuels (which often rely on food crops), advanced biofuels are produced via the conversion of lignocellulosic biomass, such as inedible plant matter, energy crops, and waste.

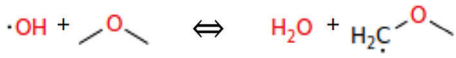

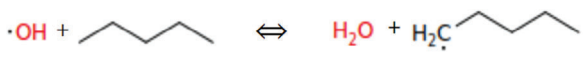

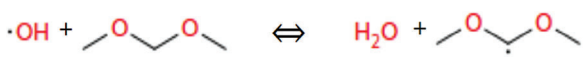
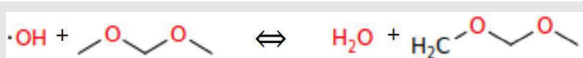
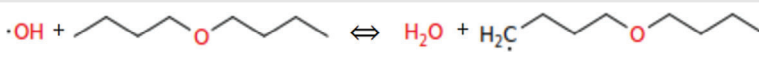
To effectively utilize the developing range of advanced biofuels available, it is necessary to determine their impacts on the performance and emissions of conventional engine technologies. Exhaustively characterizing each new fuel (and their various blends) experimentally is an extremely costly and time-consuming process, while the need to eliminate the use of fossil-derived fuels is urgent. Computer modelling provides an opportunity to predict combustion behavior relatively cheaply and quickly but requires a robust, detailed chemical kinetic mechanism, that is capable of simulating combustion behavior over wide ranges of pressures, temperatures, and equivalence ratios. The large, detailed mechanisms that typically describe the combustion of advanced biofuels may contain hundreds of species and thousands of reactions. Examples of this can be seen in the literature for detailed mechanisms describing the combustion of diethyl ether (DEE) (746 reactions, 3555 species),¹⁰ anisole (1240 species, 6004 reactions),¹¹ n-octanol (1280 species, 5337 reactions),¹² and blended biofuel mixtures such as the ethyl levulinate, ethanol, and diethyl ether blend mechanism (575 species, 1657 reactions) of Howard

et al.¹³ Creating highly accurate, detailed mechanisms by hand is often tedious, error-prone, and requires extensive expert knowledge of reaction pathways. Also, for combustion systems, the reaction rate coefficients for the vast majority of these reactions are uncertain, due to a lack of direct experimental measurements or high-level theory calculations.¹⁴ Only a small fraction of the rate coefficients of relevance for advanced biofuel oxidation mechanisms have been studied directly, and of those, even less have been studied at combustion relevant conditions. The experimental measurement of thermodynamic and transport properties for hundreds of relevant species, or the determination of thousands of rate coefficients is clearly infeasible. Accurate quantum mechanical calculations for such parameters could reduce this workload but would still be extremely computationally demanding. Approaches which can generalize information obtained from experimental measurements or theory-based calculations for well-studied species and reactions, and use this to estimate unknown reaction rate parameters and species thermochemistry, could potentially provide a solution to these issues.

1.1 | Automatic mechanism generation

Recently, the chemical engineering community has developed a contagious excitement regarding the prospect of machine learning (ML) and artificial intelligence (AI) techniques entering the research space, with the potential to assist in the development of new technologies and solve historic problems.^{15–17} Regarding the development of detailed kinetic models, it is often claimed that ML can be applied to algorithmically predict fundamental thermokinetic information (such as species thermodynamic properties and reaction rate parameters) for unexplored species, based on decades of carefully determined data for more conventional species containing similar functional groups. Effectively, data-driven predictions like these can be applied autonomously to aid in the development of detailed kinetic models through automatic mechanism generation (AMG). AMG methods also limit the likelihood of human error by automating time consuming processes and embedding expert understanding.¹⁸

TABLE 1 An example showing how training data for small species and well characterized reactions may be utilized in a reaction family to predict reaction rates for larger species with similar structures.

	Reaction	Rate coefficient ($k(T)$) ($\text{m}^3/(\text{mol}\cdot\text{s})$)
Training reaction #1		$9.35 \times 10^{-1} T^{2.29} \exp\left(\frac{3266.45 \text{ J/mol}}{RT}\right)^{23}$
Training reaction #2		$2.26 \times 10^{-3} T^{2.93} \exp\left(\frac{16903.36 \text{ J/mol}}{RT}\right)^{24}$
Training reaction #3		$2.73 \times 10^1 T^{1.81} \exp\left(\frac{3633.39 \text{ J/mol}}{RT}\right)^{23}$
Unknown reaction #1		$2.26 \times 10^{-3} T^{2.93} \exp\left(\frac{16903.36 \text{ J/mol}}{RT}\right)$
Unknown reaction #2		$1.13 \times 10^{-3} T^{2.93} \exp\left(\frac{16903.36 \text{ J/mol}}{RT}\right)$
Unknown reaction #3		$9.35 \times 10^{-1} T^{2.29} \exp\left(\frac{3266.45 \text{ J/mol}}{RT}\right)$
Unknown reaction #4		$2.73 \times 10^1 T^{1.81} \exp\left(\frac{3633.39 \text{ J/mol}}{RT}\right)$

The potential of the AMG concept has inspired the creation of several tools, featuring a variety of implementation methodologies. However, the basic underlying principles of AMG are common throughout. All possible reactions which may occur between a set of user-defined reactants, are predicted by a database of reaction ‘families’. These families may describe a reaction template and provide rate estimates based on large sets of training data, for reactions which display similar structural changes. Examples of such families include hydrogen abstraction, internal hydrogen migration, and cyclic ether formation. In this way, training data for the reactions of well characterized species can be applied to estimate reaction rates for species which have not been extensively studied. A selection of simple examples for this process can be seen in Table 1. In this example, several unknown hydrogen abstraction reaction rates for oxymethylene ether (OME₁) and di-n-butyl ether (DNBE) are estimated based on ‘training’ reaction rates for the relatively well-studied species of dimethyl ether (DME), DEE, and n-pentane. In the case of unknown reactions #1 and #3, rates are estimated to be equal to those of training reactions #2 and #1, respectively. In reality, while molecular structures local to the reaction site are similar between these unknown and training reactions, these estimations will have a limited degree of accuracy, as species specific phenomena and non-nearest neighbor interactions influence the rate of reaction. This introduces uncertainty into this estimation process which may need to be addressed by the AMG user (by sourcing additional rate

data from the literature) or the AMG tool itself. An example of how this issue may be addressed during the AMG process can be seen in the calculation of rate parameters for unknown reaction #2. In this case, the rate parameters of training reaction #2 are utilized. However, the clear difference between the two relevant structures necessitates significant corrections to the training rate. For the purpose of this example, the reaction path degeneracy method is applied (as employed by RMG¹⁹) to modify the pre-exponential A-factor. Here, the reaction path degeneracy refers to the total change in spin multiplicity due to the reaction. Due to the differences in local structures, this corrected estimate may still not be sufficient and would certainly be a target for model development and refinement. Unknown reaction #4 shows a scenario for DNBE in which the nominal reaction site (the terminal carbon) is far from the ether moieties present in DME and DEE, so the use of training data for these species is inappropriate. Instead, in a case like this, it may be sufficient to appropriate the rates of straight-chain alkanes, such as the n-pentane reaction shown in training reaction #3. The influence of the ether group on local bond dissociation energies is minimal at this abstraction site.²⁰ However, site selectivity is not only due to bond dissociation energies (particularly in the case of hydrogen abstractions by OH radicals), and interactions between the abstracting radical and oxygenated functional groups can cause some sites to be favored. An example of this can be seen in straight chain alcohols such as propanol or n-butanol, in

which the O atom of the alcohol group forms a stabilizing hydrogen bond with the abstracting OH radical, facilitating abstraction at the γ site.²¹ A similar phenomena is also observed for hydrogen abstractions from esters.²² Estimating such interactions is beyond the capabilities of these simple rate estimation methods, highlighting the need to embed accurate, relevant, and up-to-date rate parameters in the estimation process where possible.

Some AMG tools (such as MAMOX)²⁵ incorporate reaction lumping to minimize the amount of required reactions and thus, the size of the generated mechanism, whereas others will generate elementary reactions only. As the kinetic model is enlarged algorithmically, through the exploration of reaction pathways and generation of rate parameters, the AMG tool must determine which reactions and species are necessary. Termination criteria are implemented which aim to include only important reactions and limit the size of the generated mechanism. These criteria will vary between different AMG tools. Early termination methods utilized a rank-based method, limiting the maximum number of reaction steps.^{26,27} Tools such as Genesys²⁸ employ a rule-based termination criteria, which employs pre-defined constraints for each reaction family. Reactions are only included in the mechanism if all these constraints (or rules) are met. These rules may require user-definition for each reaction family, requiring knowledge of chemical kinetics. Rate-based termination criteria rely on reactor simulations to determine which species and reactions should be included in the mechanism, eliminating reactions and species which do not meet minimum reaction flux requirements. As this method is simulation based, it relies heavily on accurate species thermochemistry and reaction rate parameters. A rate-based method is employed by RMG,²⁹ the description of which is expanded upon in Section 2.1.

During the process of reaction exploration, rate parameters may be scaled based on the relative molecular structures or reaction degeneracy.¹⁹ For reversible reactions, reverse rate coefficients are calculated based on the forward and equilibrium rate coefficients, maintaining thermodynamic consistency, as shown by Equation (1). Here, T is the temperature at which the reaction is occurring, R is the molar gas constant, ΔG_{rxn}^0 is the standard free energy of the reaction, Δn is the change in moles, P is the standard pressure, K_{eq} is the equilibrium constant, and k_f and k_r are the forward and reverse rate coefficients, respectively.

$$\frac{k_f}{k_r} = K_{eq} = \left(\frac{RT}{P}\right)^{-\Delta n} \exp\left(\frac{-\Delta G_{rxn}^0(T)}{RT}\right) \quad (1)$$

Species thermochemistry is critical to the determination of reverse reaction rates. If known, this data may be supplied by the user, in NASA polynomial format.

For species where no experimentally determined or quantum mechanically calculated data exists, thermochemical parameters are often estimated using Benson-style group additivity (GA) methods.³⁰ This method assumes that the thermochemistry of a given molecule can be calculated via the contributions of all the various 'groups' within the molecule. For this method, a group is defined as a central polyvalent atom and all neighboring atoms. The implementation of the group additivity method may vary between different AMG tools, but values for stable molecules are typically calculated as the sum of group contributions, determined from the experimental measurement or quantum mechanical calculation of molecules which share groups. These values have been collated and summarized multiple times in the literature and are constantly subject to updates, as well as the definition of new groups.^{31–35} These calculations may be subject to various correction factors, depending on the species, including non-next nearest neighbor interactions (NNI), ring strain, and resonance corrections.^{31,32,35} Group additivity estimates for free radicals are also subject to a 'correction', as individual group values are understandably difficult to determine for such species. The hydrogen bond increment (HBI) method is commonly applied to calculate estimates for radicals (R) based upon the thermodynamic parameters of the parent molecule (RH) and a correction to account for the loss of a hydrogen atom.³⁶ The accuracy of this method is highly variable and is strongly dependent on the availability and accuracy of data for individual groups, but when solved algorithmically, the calculation method is rapid and can produce enthalpy of formation estimates within 5 kJ/mol of the nominal values.^{31–35} However, combustion simulations of key targets have been shown to display significant sensitivity to species thermochemistry and individual group contributions, so even small uncertainties in this method can propagate through the model.^{37,38} Therefore, it is necessary that tools include a comprehensive, up-to-date database of group contributions and corrections, which can be easily modified by users.

To distinguish between individual species and enable the algorithmic identification of structural groups, each species requires an unambiguous, unique species representation. This is not a trivial process considering the potential size and complexity of detailed mechanisms. Species must be defined such that no duplicates can exist, and the molecular structure is described as completely as possible. This representation can then be used to search for functional groups, reaction moieties of relevance to different reaction families, and GA contributions, including the necessity for GA corrections. The estimation of unknown species transport properties may also be dependent on group contribution methods, which

require a robust species representation. Such information is required for the simulation of reacting gas-phase flows, through the description of parameters such as diffusivity, heat conductivity, and mixture viscosity. Diffusion coefficients for each species are calculated using Lennard-Jones parameters, which are commonly estimated using group contribution methods, such as the Joback GA method.^{39,40} It is also not uncommon for estimates of unknown transport parameters to be made by empirical correlations, when this data are not provided by the user.^{29,41}

Various AMG strategies have been extensively described and reviewed previously in the literature.^{14,42–44} Commonly utilized AMG tools include Genesys (Ghent University),²⁸ EXGAS (CNRS and the Université de Nancy),⁴⁵ MAMOX (Politecnico di Milano),²⁵ KUCRS (University of Tokyo),⁴⁶ and RMG (Massachusetts Institute of Technology).²⁹

De Bruycker et al.⁴⁷ applied Genesys to generate pre-nol and iso-pre-nol sub-mechanisms for the development of an oxidation and pyrolysis model. For these automatically generated sub-mechanisms, elementary reactions were classified into various reaction families. These families were used to determine rate coefficients for most reactions, with rates estimated by various techniques (including group additivity, analogy to known reactions, Evans-Polanyi, and reactivity structure-based rate rules). The initial model produced by De Bruycker et al.⁴⁷ severely underpredicted acetaldehyde concentrations measured in a jet stirred reactor (JSR), but this was improved significantly with updates based on reaction path analysis. Hakka et al.⁴⁸ produced a kinetic model for predicting the oxidation of methyl and ethyl butanoates, using EXGAS. This model was then evaluated against experimental JSR concentration measurements and shock tube ignition delay times (IDTs). Ultimately, the study found that their model was incapable of correctly predicting the conversion of methyl butanoate and the formation of ethylene (with respect to JSR measurements). The underlying chemistry driving the oxidation was investigated by reaction flux and sensitivity analyses, which identified ethyl esters as a potentially significant source of carboxylic acid, limiting their viability as an alternative fuel component. Sakai et al.⁴⁹ produced a kinetic model for the prediction of DEE autoignition, using a combination of quantum chemistry calculations and KUCRS rate rules. Relevant rate rules were derived from works on the oxidation of iso-octane and butanol isomers^{50,51} and applied to DEE hydrogen abstractions. The DEE model was evaluated against shock tube (ST) and rapid compression machine (RCM) IDTs, and it was shown to reproduce the experimental data well at the majority of the conditions examined.

RMG is one of the most frequently utilized AMG packages, commonly applied in the construction of detailed

kinetic models, as well as for the estimation of thermochemical properties.^{35,52} As such, a literature review of the numerous instances in which RMG has been applied would be beyond the scope of this work, though some studies of interest are summarized here. Early applications of RMG were understandably focused on generating models for simple alkanes. One example of this can be seen in the study of Van Geem et al.,⁵³ which generated a pressure dependent model for the steam cracking of n-hexane. This model was capable of accurately predicting the conversion of n-hexane and the yields of major products, without fitting any kinetic parameters to their experiments. A substantially reduced version of the model managed to retain a good degree of agreement with experimental data while improving computational time, indicating that the large models initially produced by RMG may not always provide the most efficient modelling solution. The combustion of alcohols was another early target for RMG studies, likely due to the significant research interest in these species as alternative fuels. In 2011, an RMG generated n-butanol pyrolysis model was published by Harper et al.^{54,55} and compared to newly acquired pyrolysis data. Reaction flux and sensitivity analyses highlighted important reactions, which were targeted for model refinement via quantum chemistry calculations. The final model was further evaluated against a wide range of literature data including JSR measurements and shock tube IDTs, performing well throughout.

As the development of RMG continued and the community became more familiar with AMG techniques and applications, the modelling of more complex systems with RMG generated mechanisms became more feasible. By 2014, Allen et al.⁵⁶ had produced an RMG generated model, using the now deprecated RMG-Java (version 3.3),⁵⁷ for investigating the combustion chemistry of di-isopropyl ketone (an analogy for a series of potential advanced bio-fuels produced by endophytic fungi). For the mechanism generation, Allen et al.⁵⁶ used built-in RMG libraries to supply thermodynamic data and to seed (a concept which is discussed in the Section 2.1 of this work) the small species core of the reaction mechanism. When initially evaluated against pyrolysis and oxidation speciation data, as well as RCM IDTs, the model struggled to replicate the NTC (negative temperature coefficient) region intensity present in the experimental results. However, the authors then identified key species and reactions through sensitivity analysis, making refinements to the mechanisms based on quantum chemistry calculations. These refinements led to a significant improvement in model predictions and identified the importance of accounting for resonance stabilization in the estimation of species thermochemistry.

Recently, Dong et al.⁵⁸ (2023) investigated the combustion and pyrolysis chemistry of butyl acetate, using newly

generated RMG mechanisms. Accurate parameters for species and reactions were calculated prior to mechanism generation and stored in the RMG database, allowing RMG to utilize them automatically. The identification of influential species via flux analysis and the impact of standard enthalpy of formation ($\Delta H_{f,298K}$) perturbations, allowed for model improvement through the iterative refinement of uncertain thermochemistry. This produced a model with good predictive capabilities. However, the authors identified that a major flaw of the RMG generated model could be found in the uncertainty of sensitive low temperature oxidation pathways, due to a lack of accurate training data and a reliance on highly uncertain rate rules. This caused a limited degree of accuracy in model predictions, particularly at low temperatures.

It is clear from this selection of literature that RMG (and AMG as a whole) is a powerful tool, with the potential to produce highly accurate, robust models when applied correctly. However, in its current state, the process is not completely autonomous and can require extensive input by a user with expert knowledge to achieve the desired goal. This is not necessarily a negative; a minimum knowledge requirement will always be necessary to triage issues, provide expert critical analysis, and prevent the propagation of incorrect information. However, the time investment required to become proficient with AMG tools (such as RMG) and to develop a reliable mechanism generation methodology can be comparable to that of producing a detailed mechanism by hand, eliminating one of the major advantages of AMG. On the other hand, once the tools and training data have been established, they can be more routinely applied to other molecules with similar functional groups, and thus for developing mechanisms for complex blends, AMG methods could offer substantial savings. In this study, a general methodology is proposed for the generation of mechanisms describing the combustion of advanced oxygenated biofuels using RMG, based on the lessons learnt from the literature and experience. As an example, a mechanism is generated for the prediction of DEE oxidation and compared against various literature mechanisms and experimental data. The effectiveness of RMG when utilized for this purpose is evaluated and recommendations are made regarding the application and development of AMG tools.

1.2 | Diethyl ether

DEE is selected for this case study due to its nature as an oxygenated hydrocarbon, which has received significant research attention during recent decades as a potential biofuel. This is largely due to its ease of production, high cetane number and preferential physical properties rela-

tive to similar species, such as dimethyl ether (DME).^{59–62} As an advanced biofuel, DEE may be produced directly from cellulosic ethanol via an acid catalysed dehydration⁵⁹ or as a by-product from the ethanolysis of lignocellulosic biomass.^{63,64} DEE is highly reactive and volatile (with a cetane number of approximately 139, boiling point of 35°C, and a flash point of -45°C ⁶⁵), limiting its use as a fuel in its pure form. However, DEE has been extensively studied as a blending component for use in compression ignition (CI) engines, with ethanol^{65–67} and conventional diesel fuels.^{68,69} As a result of the potential of DEE as a biofuel, a large amount and variety of experimental data relative to the combustion of DEE already exists in the literature. This includes IDTs measured in RCMs and shock tubes^{67,70–73}; speciation measurements collected from JSR, plug flow reactor (PFR), and shock tube experiments^{10,74–78}; and several studies of DEE flames.^{79–84} In parallel to this experimental exploration of DEE combustion, several studies have attempted to produce detailed kinetic models to predict the observed behavior.

In 2010, Yasunaga et al.⁷² published one of the earliest chemical kinetic studies of the DEE pyrolysis and oxidation systems, including a detailed kinetic model for high-temperature ignition, validated against shock tube IDTs and speciation measurements. Sakai et al.,⁴⁹ building on the work of Yasunaga et al.,⁷² calculated species thermochemical data and reaction rate coefficients for important low-temperature oxidation reactions⁸⁵ and utilized them in the development of a DEE mechanism for low and high temperature ignition. This model produced a good agreement with IDT measurements at engine relevant conditions. However, Sakai et al.⁴⁹ also made arbitrary modifications to their derived thermochemical data, tuning the mechanism to achieve more accurate predictions (particularly in the NTC region). Detailed sensitivity and uncertainty analyses were subsequently performed on the Sakai mechanism by vom Lehn et al.³⁷. Using uncertainty minimization techniques, they produced an optimized version of the Sakai mechanism which reproduced experimental IDTs more accurately. Tran et al.¹⁰ published a new, detailed model, not based on solely the mechanisms of Yasunaga et al.⁷² and Sakai et al.,⁴⁹ in 2019. The authors were motivated to create a new mechanism by the poor representation of their JSR and PFR measurements provided by the available models at the time. Jet-stirred reactor gas chromatography (JSR-GC) measurements for the conversion of DEE and the production of various important intermediates were well represented by the model, though the model is comparatively poor at predicting IDTs, particularly in the NTC region. In 2021, Danilack et al.⁸⁶ published an investigation into the influence of non-Boltzmann reactions on the ignition behavior of DEE, using a newly generated DEE

mechanism. The authors computed stationary points on multiple potential energy surfaces of relevance to DEE oxidation, at the UCCSD(T)-F12b/cc-pVTZ-F12 level of theory.^{87,88} They then went on to calculate the corresponding temperature and pressure dependent rate constants, using the RRKM/ME code MESS and microcanonical transition state theory.^{86,89} However, this mechanism only considered a single DEE chain-branching pathway. In a following publication, Danilack et al.⁹⁰ expanded on this mechanism to include five additional chain branching pathways, to investigate the potential influence of O₂QOOH diastereomers on DEE auto-ignition. In both cases, the studies were entirely theoretical endeavors and neither mechanism was validated against experimental data. From this brief summary of detailed DEE mechanism development, it is apparent that no single mechanism is effective at accurately predicting both autoignition and speciation measurements accurately. As such, there is a need for a robust general purpose DEE combustion model.

2 | METHODOLOGY

The development of a reliable mechanism generation methodology using RMG, specifically for the prediction of oxygenated advanced biofuel combustion, is one of the core objectives of this study. Understanding the evolution of the final methodology and the lessons learnt during this process is vital to discussing the implications for the current state and future development of AMG tools. As such, it would be inappropriate to detail the final mechanism generation methodology in this section. Instead, this methodology section is dedicated to describing general RMG use and providing simulation details. Subsequent sections are then dedicated to describing and evaluating the various mechanism generation methodologies investigated in this study, their impacts on the efficacy of the resulting mechanisms, and their additional analysis and data requirements.

2.1 | Reaction mechanism generator

Full details of the RMG methodology and the theory underpinning its operation are provided in the literature.^{19,29} General AMG methodologies have been summarized in Section 1.1. The RMG mechanism generation methodology is summarized here from the perspective of a user, to ensure a foundational level of understanding regarding the mechanism generation process, and to inform the necessity of the subsequent changes for the accurate prediction of advanced biofuel combustion phenomena.

To initialize mechanism generation, RMG requires the user to provide an input file which contains specifications for various necessary parameters. This includes reactor conditions, starting species concentrations, termination criteria, and model tolerances. Important species and reactions are determined using an implementation of the Susnow et al.⁹¹ rate-based algorithm. The user-defined set of initial species (the initial model core) are reacted under the specified initial conditions to generate a list of possible product species (the model edge). A database of reaction families and training data are employed to enable this exploration of possible reactions and species, as described generally in Section 1.1. The latest stable version of the RMG database (3.0.0) contains an extensible set of 77 reaction family templates.^{19,92} However, not all of these are applicable to the oxidation and pyrolysis of gas phase mixtures. For these cases, by default RMG utilizes 51 of these reaction families, though if the user wishes to include others, these can simply be specified within the input file. While RMG features a wide range of reaction family templates, the amount of training data incorporated into the database is heavily biased towards a much smaller number of families. Approximately half of the reaction families contain less than five training reactions and only 17 families contain more than 20 reactions.⁹² These 17 reaction families contain a total of 8831 training reactions, with 3107 and 2894 of these reactions being utilized as training data for the two largest families, 'H_abstraction' and 'R_Addition_MultipleBond', respectively. To determine which reaction templates and training data are applicable for a given species, each molecule is described by a unique adjacency list: a graph data type representation of the atoms and bonds in a given molecule, the syntax of which is described in detail in the literature.²⁹ Example adjacency lists for DEE, the associated fuel radicals, and selected oxidation intermediates are presented in [Supplementary Materials](#), while a full description of the adjacency list notation can be found in the RMG literature.²⁹ Using this representation, functional groups can be identified. These groups are also described by adjacency lists, which differ slightly, in that atom types are used in the list definition, rather than specific elements. These atom types describe local bond structures and a comprehensive list is available in the literature,²⁹ so it is not repeated here. The definition of species and functional groups in this manner allows high quality training data acquired for a well-studied species to be generalized and algorithmically applied to a species with an unknown reaction rate. Reaction rate parameter predictions are automatically scaled by RMG based on reaction path degeneracy.¹⁹ For the generation of an accurate detailed mechanism, this methodology is reliant on the provision of relevant, high quality training data

(experimental measurements or low uncertainty quantum mechanical calculations), specific to the functional groups of interest. If these are not present, inappropriate reaction rates will be applied and can propagate leading to potentially large uncertainties in model predictions. An example of this would be the use of a hydrogen abstraction training rate derived from straight chain alkanes (i.e., propane) when attempting to predict a rate for the secondary hydrogen abstraction of DEE. In this case, the presence of the ether group would significantly decrease the relevant C-H bond dissociation energy, thus accelerating the associated rate of reaction.^{93,94} Appropriate training reactions are generally lacking for oxygenated species throughout the database of reaction families, particularly for ethers, esters, and molecules which feature multiple oxygenated functional groups.

The reactor simulation may run to completion or terminate early if the reaction flux to an edge species exceeds a user-specified tolerance. In each case, edge species are then brought into the model core if their flux exceeds the characteristic flux of the system.²⁹ By default, only the edge species with the largest flux is brought into the core, and the process is repeated until the global termination criteria is reached (as defined by the user in the input file). However, RMG also allows the user to specify how many species can be moved from the edge to the core in a single simulation, to enhance the tools ability to identify competing reactions pathways, like those typical of low temperature oxidation.^{22,95} The accuracy of thermochemistry and reaction rate parameters of relevance to the species of interest is critical to the generation of an appropriate model core. Ultimately, the model core will describe the contents of the generated mechanism.

Thermodynamic parameters for species may already be present in the RMG database, contained within thermodynamic libraries. These libraries catalogue high-quality thermochemical data sourced from experimental measurements and quantum mechanical calculations, which can be accessed by RMG during mechanism generation. However, the majority of these libraries focus the provision of thermochemistry for small species typical of all hydrocarbon combustion systems, such as alkanes, cyclic species, organic molecules containing nitrogen or sulfur, and the associated combustion intermediates.⁹² Little thermochemical information is provided for the complex oxygenated hydrocarbons typical of advanced biofuels. The user can create a new thermodynamic library describing the thermochemistry of their species of interest, but accurate values for low temperature oxidation intermediates are unlikely to exist in the literature for more complex or emerging biofuel species. In this case, RMG will estimate species thermochemistry using a GA methodology. This method broadly follows that described in Section 1.1,

with corrections (of relevance to gas phase kinetics) for radical species (using the HBI method), NNIs (including non-cyclic gauche interactions), ring and polycyclic strain, and ketenes.^{36,92,96–101} The group database alone (without any correction groups) contains 2085 groups and their associated contributions (RMG database version 3.0.0). However, while not necessary for the prediction of thermochemistry in this case study, this database still requires expanding and updating to increase the accuracy of group attribution and estimates for the oxidation intermediates present during the combustion of complex oxygenated biofuel species, such as ethyl levulinate. In special cases, RMG can also predict species thermochemistry using a machine learning model or on-the-fly quantum mechanical calculations, but these are not utilized in this study due to poor estimation performance for oxygenates and computational cost, respectively.^{19,29,92}

Reactor conditions directly impact the included species, reactions, and therefore the conditions at which the model is applicable. The reactor conditions for which the model is generated, must describe the thermodynamic and mixture conditions of the final intended application, in terms of temperature, pressure, and initial mole fractions. Ranged reactors are an RMG feature which allows the user to specify a range of reactor conditions (rather than a series of fixed-point reactors), which are then sampled using a weighted stochastic grid sampling algorithm. RMG documentation claims that these ranged reactors reduce the number of simulations needed during the generation process and improve the representation of reaction conditions that would otherwise occur in the thermodynamic space between individually specified reactors. This feature is also more convenient for the user, relative to the definition of potentially dozens of individual point reactors.¹⁹ For gas-phase applications, RMG can currently only model isobaric, isothermal, homogeneous batch reactors (defined as a simpleReactor in RMG).^{19,29}

For the purpose of this study, ranged reactors are defined to cover a wide range of conditions relevant for low temperature combustion. This includes temperatures of 550–1100 K, pressures of 1–40 bar, equivalence ratios (ϕ) of 0.5–2.0, and a range of diluents (Ar, CO₂, N₂). Diluents and equivalence ratios are specified by describing the starting species concentrations, within the reactor block. Reaction systems are simulated until pre-determined termination criteria are met. These criteria may be specified as a termination time, species concentration, rate ratio, or some combination of these. In this study, simulations are defined such that they terminate at a simulated time of 1 s or at a conversion of 99.9% of DEE, whichever occurs first. While not usually important at engine relevant pressures, pressure dependent reaction networks for C₄ and larger molecules are known to have a considerable influence at

the low pressures and high temperatures typical of laboratory flame speed experiments.¹⁰² For this reason, the RMG pressure dependence module is utilized to estimate pressure dependent kinetics based on a simplification of the full master equation model: the modified strong collision method.^{103,104} This method usually provides good accuracy with a much lower level of computational cost.^{29,103}

Model tolerances limit which species are included in the model core, as defined by the user specification of flux tolerances. A core flux tolerance which is too high may cause RMG to miss critical species and reactions, whereas a low tolerance will increase the computational cost of the mechanism generation, in terms of time for convergence and RAM (random-access memory) usage. Species and reaction data are stored in RAM during the mechanism production process, and this issue of high RAM usage has been a persistent problem for RMG.^{29,105} The issue is exacerbated in the region of low temperature oxidation, particularly for large species and complex mixtures (which are of interest for many advanced biofuels). The termination of a mechanism generation process due to exceedance of a computer's RAM limit is not an uncommon occurrence. This can be minimized by carefully creating input files, but this requires experience and knowledge of the RMG mechanism generation process and the influence of model tolerances. System specific features, such as the utilization of 'swap spaces' in modern Linux systems, can also limit the occurrence of premature model termination by artificially increasing the memory available. However, this is a persistent RMG issue and if tools such as this are to be applied for the detailed investigation of complex fuel mixtures, development beyond the exclusive use of RAM for data storage is likely required (i.e., the use of temporary files stored in memory, for information which does not need to be accessed rapidly). The mechanisms generated in this study for DEE are unlikely to meet the RAM limit of a conventional, modern system, reaching a maximum reported memory usage of 4818 MB and less than 4 h of run time. However, it should be noted that this is the memory use reported by RMG at the end of each generation step and not the peak system memory usage, which can spike considerably above this value during the generation of pressure dependent rate parameters. Such unreported memory spikes can cause issues during the mechanism generation process, including premature termination (as discussed in Section 3.4).

Even with low model tolerances, the flux criteria method employed by default in RMG is not always capable of identifying highly important but low flux pathways. This is particularly true for those pathways which result in a net increase of radicals, such as chain branching. RMG allows the user to apply specific branching criteria tolerances, which can assist the model in identifying important

edge reactions based on their expected impacts on core species concentrations. However, it is usually much more effective to utilize good seed mechanisms to fulfil this purpose. Seed mechanisms contain reaction rate information that the user wishes to include in the model core, overriding RMG's native rate estimates. RMG will then build on the model core in the usual manner.²⁹ A good seed mechanism should utilize high quality data to guide the mechanism generation process, particularly when the RMG kinetic database is insufficient. In general, it is good practice to provide a fuel specific seed mechanism which (at a minimum) describes the key initiation and branching reaction steps, as well as any low uncertainty rate data which the user wishes to include in the model. Seed mechanisms can also be utilized to embed well-defined sub-mechanisms in the initial model core. Reaction libraries are created identically to seed mechanisms but defined differently within the input file. Information from reaction libraries is not automatically included in the model core. Instead, libraries serve as a database of specific reaction data which is called upon by RMG when the reactions in question are deemed relevant. Similarly, thermodynamic information for species may be defined in thermodynamic libraries, which will override any RMG estimates. All the models generated in this study consist of a small molecule core defined by pre-packaged RMG seed mechanisms: 'BurkeH2O2inN2', 'BurkeH2O2inArHe', and 'C2H4+O_Klipp2017'.^{106,107} A C0-C4 library is also utilized, based on the extensively validated AramcoMech 2.0,¹⁰⁸ to ensure well-defined small molecule chemistry. Fuel specific seed mechanism information for each model is described in the appropriate following sections. Transport data are provided by the pre-packaged 'NOx2018'¹⁰⁹ library where possible, or otherwise estimated by RMG. All tolerances and generator options used in the production of the final DEE model are shown within the RMG input file, presented in [Supplementary Materials](#).

2.2 | Reactor simulations

The final mechanism generated in this study is evaluated against a range of DEE combustion data, sourced from various experimental studies. This includes RCM and ST IDTs,^{67,70,72} JSR-GC speciation measurements,^{10,110} and flame speeds.⁸⁰ Simulations of each of these systems are produced using Chemkin-Pro (2022 R1), using the closed homogeneous batch reactor, perfectly stirred reactor, and pre-mixed laminar flame-speed calculation modules, respectively. All simulations are conducted under the conditions specified in the original studies. For flame simulations, curvature, and slope refinement criteria of 0.05 are found to produce grid independent results, without

requiring an excessively long compute time. The predictive capabilities of the RMG mechanism generated in this study are also compared against those of three independently generated DEE mechanisms from the literature: those from Sakai et al.,⁴⁹ Tran et al.,¹⁰ and Danilack et al.⁹⁰ In the case of the Danilack et al.⁹⁰ mechanism, only the version without diastereomers included is used.

To account for RCM facility effects and heat losses after compression, case dependent volume histories derived from non-reactive pressure profiles (as provided by the original study of Issayev et al.⁶⁷) are applied to produce 'variable volume' reactor simulations. ST measurements often exhibit a gradual pressure rise behind reflected shock waves, which can be observed in experimental pressure traces. In this study, ST simulations account for this by imposing an experimentally determined pressure gradient, as defined in the data sources.^{67,70} This is not particularly important at conditions where ignition occurs rapidly but can be influential for relatively long IDTs (i.e., IDTs > 2 ms).

The final steps of the mechanism production methodology applied in this study ('post-processing'), employ model analysis techniques in the form of local A-factor sensitivity analysis for reactions, brute force enthalpy of formation sensitivity analysis, and rate of production (RoP) analysis. Local A-factor sensitivity and RoP analyses are conducted using Chemkin-Pro. Local A-factor sensitivity analysis is performed for peak OH concentrations, in the case of IDT simulations, and end-point concentrations of DEE, EVE, CH₂O, C₂H₄, C₃H₆, C₂H₅OH, and C₄H₈O₂-CY (2-Methyl-1,3-dioxolane) during the simulation of the JSR-GC experiments performed by Tran et al.¹⁰ The brute force enthalpy of formation sensitivity analysis is performed for homogeneous IDT predictions and completed using Cantera, and has been documented in detail previously.¹¹¹ In this study, these techniques are not applied to fully investigate the underlying DEE chemistry driving combustion. Instead, they are utilized to identify key reactions and species for further investigation, to improve the accuracy of the model. Where possible, highly sensitive species and reactions which also contain uncertain thermodynamic or rate parameters are updated with less uncertain rates, sourced from calculations or experimental data within the existing literature. Similarly, RoP analysis is employed to identify potentially missing or erroneous reactions. Erroneous reactions may be present in the generated mechanism due to the allocation of uncertain rates or inappropriate training data during model core enlargement steps. The case is similar for missing reactions. However, identification of missing reactions is non-trivial (even with the assistance of RoP analysis) and is often reliant on the user's knowledge of reaction pathways specific to the relevant species or functional group. The

search for missing reactions may also be informed by the automated exploration of potential energy surfaces using tools such as KinBot,¹¹² but such a search is not performed in this work. Missing reactions may also occur as a consequence of competing erroneous reactions, which (due to their exaggerated reaction flux) direct the exploration of the reaction network towards unrealistic or unimportant reaction pathways.

3 | MECHANISM GENERATION PROCESS

To determine the effectiveness of RMG for producing detailed kinetic mechanisms of complex low temperature oxidation for oxygenated fuels and to develop an optimal mechanism production methodology for such cases, DEE is used as a case study. Several development steps are investigated in the following sections, starting with the most basic use of RMG with no updates in the core training data or group specifications and without the use of a fuel specific seed mechanism. This is referred to as development stage (1). In subsequent development steps an appropriate seed mechanism is included, and updates are made to the core databases within RMG. These methods and their impact on IDT predictions are detailed by development stages (2) and (3), respectively. The impact of each method is investigated independently in the corresponding development stages. In the final stages, various post processing tools are applied to identify key species and reaction steps requiring improvements in their parameterization, as described in development step (4). A simple block diagram is provided in Figure 1 to provide additional clarity as to the structure of this description of the mechanism development. Figure 2 shows the predicted IDTs for the mechanisms produced during each development stage, compared against the experimental RCM and shock-tube data of Issayev et al.,⁶⁷ and the shock-tube measurements of Uygun,⁷⁰ for an undiluted stoichiometric DEE/oxidizer mixture. Comparative analysis of the models shown and detailed descriptions of the evolution of the model development process are provided in the subsequent sections. All newly produced seed mechanisms, reaction libraries, thermochemistry libraries, and database expansions are available in [Supplementary Materials](#).

3.1 | Development stage (1) – Minimal (RMG baseline)

The minimal method used in development stage (1), is performed only to provide a baseline 'worst-case' DEE mechanism for comparative purposes. The method described here is not advised anywhere in the RMG

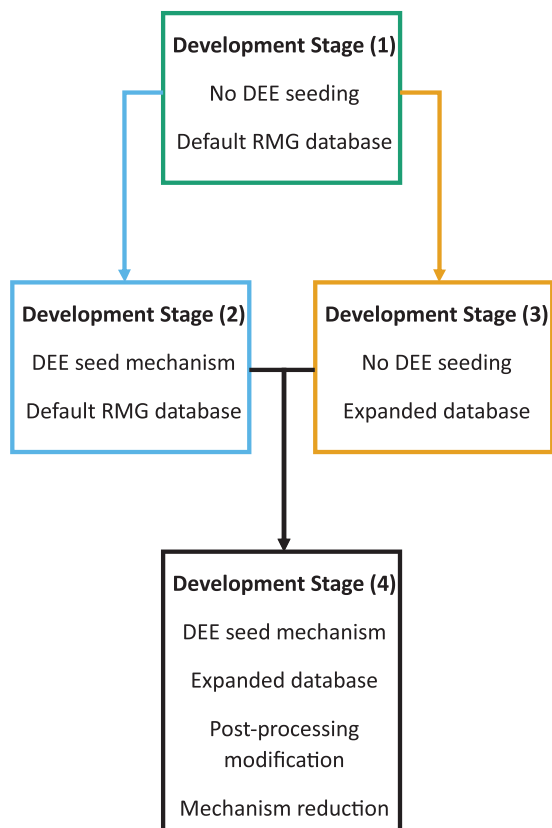


FIGURE 1 A simple block diagram to aid in the description of the mechanism development process. Refer to Section 3 of the main manuscript for detailed descriptions of each development stage.

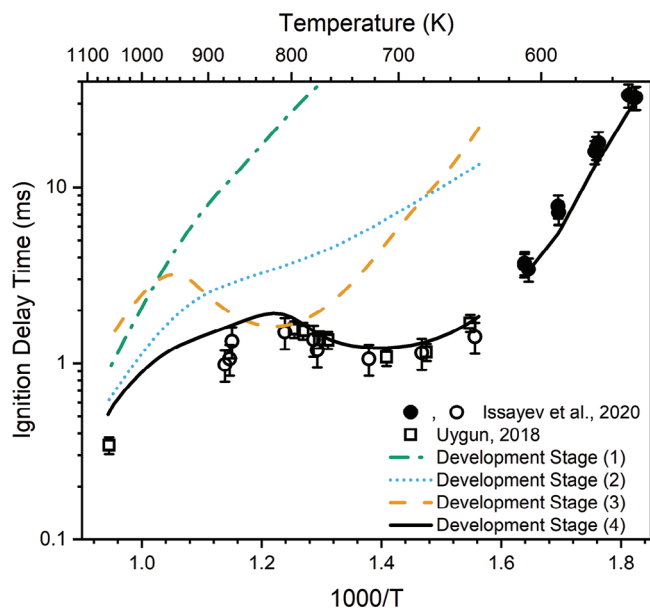


FIGURE 2 Variable volume IDT predictions of the various DEE mechanism iterations produced in this study. Symbols show the experimental RCM and shock-tube data of Issayev et al.,⁶⁷ and the shock-tube measurements of Uygun,⁷⁰ lines show the model predictions. Closed symbols represent RCM measurements and open symbols show ST measurements. Reactor conditions are $T_c = 550\text{--}1100\text{ K}$, $P_c = 20\text{ bar}$, $\varphi = 1.0$.

documentation or literature, nor is it a suitable method for the production of a detailed kinetic mechanism. However, ultimately AMG tools should be aiming to require as little user intervention as possible to truly facilitate the rapid, autonomous production of kinetic mechanisms. This mechanism is generated using a clean install of the latest stable version of RMG (RMG 3.1.0) and the associated RMG database. The mechanism generation process is only provided with the small molecule core (described previously) and basic thermodynamic libraries. These thermodynamic libraries include parameters derived from the work of Burke et al.¹⁰⁶ for the high-pressure combustion of H_2/O_2 mixtures ('BurkeH2O2'), high-accuracy ab initio thermochemistry for various small molecules as calculated by Goldsmith et al.¹¹³ ('DFT_QCI_thermo'), and the 'primaryThermoLibrary'. Initial species (fuel, oxidizer, and diluents) are specified by their canonical SMILES strings, but no other fuel specific information is provided.

This minimal method primarily shows the results of mechanism generation using RMGs default database, as all fuel specific rates are estimated using RMGs internal reaction family templates and training data. As evidenced by Figure 2, this methodology produces an unacceptably poor model. Predicted IDTs are orders of magnitude larger than the literature sourced experimental data (for the same conditions) and none of the low temperature behavior expected for DEE is observed (i.e., the NTC region). It is clear from these results that a default installation of RMG is insufficient for identifying and estimating the reaction rates of key low temperature combustion pathways in this case. The thermodynamic parameter estimations produced by RMG are also likely contributing to the poor performance of the model. Group-additivity based estimations are subject to significant uncertainties when compared to ab initio calculations or experimental measurements. The recent literature has reiterated the significant influence of uncertainties in species thermodynamic properties on the accuracy of global predictions (such as IDTs), especially in the low to intermediate temperature combustion regions.^{37,38,111,114} Uncertain fundamental data in the RMG database may further influence the mechanism generation process through the exclusion of important reactions and species (mechanism truncation error), as the calculated fluxes are not representative of the real scenario. Therefore, the specification of accurate species properties is critical to the production of a detailed model. These issues can mostly be addressed by two techniques: the provision of a high-quality seed mechanism and thermodynamic library (as described in Section 2.1), and the expansion and updating of the RMG database to include kinetic information relevant to the fuel species (and its important oxidation intermediates). Development stages (2) and (3) apply each of these methods, respectively.

3.2 | Development stage (2) – Fuel specific seed mechanism

As described in Section 2.1, the provision of a high-quality seed mechanism can aid RMG in the identification of important reaction pathways, which the flux-criteria algorithm would typically fail to include in the model core. This is predominantly an issue for the generation of low temperature oxidation mechanisms. As well as guiding RMG's reaction pathway exploration, the inclusion of a seed mechanism provides an opportunity to specify any low-uncertainty fuel specific rate parameters that the user wishes to include in the final mechanism. Similarly, thermodynamic libraries are used to specify the thermodynamic properties of species, overriding and group-additivity derived parameters.

Care must be taken when producing a seed mechanism for use with RMG, particularly when this seed is based on a literature sourced sub-mechanism. Non-elementary reactions are common even in detailed mechanisms, particularly in the representation of low-temperature degenerate chain branching steps. Several elementary reaction steps are occasionally lumped together using quasi-steady state assumption (QSSA) methods, which assume that rapidly reacting species are in local equilibrium with slower reacting species. Examples of this can be observed in the dibutyl ether mechanism of Thion et al.,²⁰ in which the decomposition of keto-hydroperoxides lumps together multiple fast decomposition reactions to give: 'C₄OC₄KETA-3 → nC₃H₇ + CO₂ + C₂H₄ + CH₃CHO + OH'. For an RMG seed mechanism, such reactions should be carefully removed from the scheme where possible, since the algorithmic exploration of reaction pathways does not distinguish between elementary reactions determined using the internal database and the lumped reactions present in the seed mechanism, resulting in the presence of duplicate reaction paths within the final scheme. While heavy thermodynamic library loading (the inclusion of many thermodynamic libraries) is a common and perfectly acceptable practice when applied carefully, the same techniques should not be applied to the specification of seed mechanisms within the RMG input file. Not only will this cause the generation of an unnecessarily large mechanism, but the use of seed mechanisms which are unconnected to the initial model can cause solver issues, as simulations struggle to converge. An unconnected mechanism in this sense refers to a mechanism which is unrelated to the system of interest (but perhaps was included by the user as it contains some rate parameters of relevance) or describes reaction pathways which cannot be connected to those present in the initial core by the addition of RMG's elementary reaction families. Using unconnected

seed mechanisms as reaction libraries is an appropriate alternative.

Development stage (2) incorporates a DEE seed mechanism derived primarily from the DEE sub-mechanism of Tran et al.,¹⁰ containing 86 reactions. Key reactions covered by the seed include initiation H-abstractions of DEE by H, OH, and HO₂ radicals, primary, secondary, and tertiary oxygen additions, subsequent internal isomerization reactions, HO₂ elimination, and cyclic ether formation. All the rate parameters present in the seed are derived from experimental measurements or detailed calculations. Reaction rates based on estimations or correlations with other species are not included in the seed mechanism, and are generated by RMG, using the unmodified RMG database (v3.0.0).⁹² The DEE mechanism of Tran et al.¹⁰ is chosen as the basis for this seed due to the model's extremely good agreement with experimental JSR-GC data, despite relatively poor IDT predictions when compared to other DEE mechanisms in the literature.^{37,49,67}

The conversion of literature sourced data to an RMG seed mechanism is not always an easy process, despite the availability of a pre-packaged RMG tool which converts Chemkin format mechanisms into RMG reaction and thermodynamic libraries. The opportunity for human error occurs as a result of the requirements of the conversion tool and the often-inconsistent species notation present in literature sourced mechanisms. To parse the Chemkin format inputs, the RMG tool requires the specification of a species dictionary which describes each species in the mechanism by its adjacency list. Adjacency lists can be determined by tools available on the RMG website, if the species identifier is known (supported formats include SMILES, InChI, CAS number, and some species names). Fortunately, the majority of relevant species in the Tran et al.¹⁰ mechanism are provided with a SMILES string by the original authors, which removes the scope for human error due to misinterpretation of the species notation. However, the by-hand determination of species adjacency lists from this information and the production of an RMG format species dictionary is still a time-consuming process, particularly for larger mechanisms (i.e., AramcoMech 2.0). Species thermodynamic data prescribed by CBS-QB3 level calculations in the Tran et al.¹⁰ mechanism is extracted and converted to a new RMG thermodynamic library and included in the mechanism generation process. Species thermochemistry derived from GA methods is not extracted from the Tran et al.¹⁰ mechanism, as this will be generated by RMG.

As shown by Figure 2, the introduction of species-specific reaction rate and thermodynamic data, through the inclusion of a seed mechanism and thermodynamic library in development stage (2), displays an obvious

TABLE 2 A summary of the number of training reactions and structural groups added per reaction family, during the database expansion.

Reaction family	Training reactions added	Structural groups added
Cyclic_Ether_Formation	57	6
Disproportionation	4	0
H_Abstraction	349	17
HO ₂ _Elimination _from_PeroxyRadical	11	2
intra_H_migration	40	10
R_Recombination	3	7
R_Addition_MultipleBond	6	1

Note: Reaction families are identified by their RMG assigned names.

improvement in the accuracy of DEE IDT predictions relative to the minimal mechanism generated in development stage (1). In the seeded model, IDT predictions are much shorter throughout the investigated regime, with some evidence of an NTC-like region starting to develop. This region is not strictly an NTC region, but it does show a locally significant decrease in reactivity with increasing temperature. IDT predictions at the high-temperature end of the regime (>1000 K) appear to be reasonably in line with experimental measurements, although there is a lack of experimental data in this region to evaluate against. There is clearly still scope for major improvements in the model's predictive capability, as apparent in the lack of NTC intensity and the large under-prediction of low temperature reactivity.

3.3 | Development stage (3) – A database expansion for oxygenated functional groups

The lack of appropriate training data for various oxygenated functional groups (i.e., ethers, esters, ketones) in the default database leads to the inclusion of inaccurate rate data in the model, as derived from training data for inappropriate molecular structures. Development stage (3) aims to minimize the misappropriation of training data by expanding the internal database with contemporary reaction rate parameters for the oxygenated functional groups typical of biofuels. This database expansion includes the addition of 470 training reactions and 43 structural groups, primarily sourced from the literature,^{10,49,85,115–137} with a minority of the new rates provided by CBSQB3 level calculations available in the RMG GitHub repository. Table 2 presents a summary of these expansions, showing the number of reactions and groups added per reaction family. Care must be taken when performing such a database

expansion, with priority given to low uncertainty, high level of theory calculations and experimentally determined rate parameters. Modified database files are provided in the [Supplementary Materials](#), which detail the source for each additional training reaction.

For new or under-represented functional groups, new training data additions should focus initially on fundamental examples of the group of interest, before building the molecular chain length until it is representative of the nominal species. For example, several new hydrogen abstraction training reactions are included for dimethyl ether (DME),¹²⁷ wherein the carbon chain length on either side of the ether group is only one carbon. The chain length is then increased to the two carbons present in DEE, by utilizing rate parameters for ethyl methyl ether (EME).¹²⁷ For species which may feature longer carbon chains connected to an ether or ester group (if no specific information is apparent for relevant esters), such as butyl levulinate, the chain length of the training species is extended to cover methyl propyl ether,¹¹⁵ which features a three carbon long chain on one side of the ether group. Beyond this chain length the ether group is unlikely to significantly contribute to the bond dissociation energy of the terminal C-H bonds, so such a H abstraction reaction can be estimated using rate parameters derived for alkanes. However, care must be taken to ensure that such sites are not subject to effects of stabilizing hydrogen bonding between the oxygenated functional group and abstracting radical, as mentioned in Section 1.1. New structural group definitions are required to accurately describe the functional groups and reaction sites present in the additional training data. Groups are defined by RMG adjacency lists, as previously described. Unfortunately, no pre-packaged tools are readily available for the expansion of structural group information for reaction families, so these must currently be determined and carefully updated by-hand, to avoid introducing potential sources of error and uncertainty. The over-specification of groups, wherein the group is so rigorously defined that it cannot be applied to other similar species, is a potential pitfall in this process that the user should be aware of. An automated tool for the identification and addition of new structural groups for database expansion could reduce the likelihood of such errors and ultimately improve the mechanism generation process for complex species and mixtures, such as those present in advanced biofuels. Database updates in general can be difficult and time consuming to perform. A pre-packaged Python script is available to convert RMG libraries into training data, but the limited documentation available for some of the standalone tools packaged with RMG can make the process inaccessible and intimidating for inexperienced users and limits the effectiveness of mechanism generation.

After completing this database expansion, the mechanism is generated repeating the process applied in development stage (1) (i.e., no DEE seed mechanism is used). Despite the identical mechanism generation process, the mechanism produced in stage (3) shows a clear improvement over the predictions of that produced during stage (1), without the need to include a DEE specific seed mechanism. In particular, the presence of an NTC region is an obvious improvement over the previous model, which did not feature any of the expected low temperature oxidation behavior. This highlights the importance of accurate training data for generating models capable of predicting complex reaction pathways. In this case, the additional training data have allowed RMG to capture some of the behavior present in the low temperature oxidation of DEE, which relies on a complex balance of competing chain terminating and branching pathways. The expanded database utilized in this development stage is also applicable for the generation of further oxygenated species in the future, without the need to repeat any of the database expansion performed here.

While the improvement in predictions relative to the mechanism produced in stage (1) is dramatic, there are still issues with the mechanism generated in stage (3), which are apparent in Figure 2. Whilst the NTC behavior predicted by the model developed in this stage is very clear, it occurs in the incorrect temperature region for DEE. Experimental data shows that, at these conditions, the NTC region occurs at much lower temperatures, peaking at approximately 800 K rather than the predicted 950 K. IDTs are also too long throughout the regime, predicting a less reactive mixture than experiments would suggest. This would clearly be an issue for the prediction of engine performance and emissions, as the IDT is a critical parameter for modelling the developing combustion environment. Ultimately, it must be the end goal to utilize such mechanisms and their generation techniques for the rapid development of alternative fuel strategies through the rapid screening of their performance and emissions impacts with the aid of simulation tools. Therefore, it is critical that properties like IDTs are predicted accurately throughout the engine relevant regime. Evidently, for the RMG mechanism to produce accurate predictions for advanced biofuels, a combination of the methods utilized in development stages (2) and (3) is required.

3.4 | Development stage (4) – Combination and post-processing

The use of the DEE seed mechanism of stage (2) and the database expansion of stage (3) are combined in development stage (4), to produce the final exploratory DEE

mechanism. To more accurately predict the formation and consumption of ethyl vinyl ether (EVE), a small sub-mechanism is generated using the same method described in stage (1), which is then combined with the DEE mechanism generated in this section. A separate sub-mechanism is used, rather than including an additional simpleReactor block for EVE within the RMG input file for the DEE mechanism, to avoid a memory (RAM) limit issue during the generation of pressure dependent reactions. During the calculation of pressure dependent reaction rates (using the modified strong collision method), memory use increases drastically, creating a spike in usage. As the model edge and core become larger, the memory required to store relevant reaction rate and species thermochemistry information increases. Spikes in memory use can then cause the generation process to momentarily exceed the memory limit of the system, forcing the mechanism generation process to prematurely terminate. In order to preserve pressure dependence, it is necessary to limit the size of the model core and edge by generating the EVE sub-mechanism separately. During DEE oxidation, EVE is primarily formed via HO₂ elimination and hydrogen abstractions from fuel radicals. The relationship between this intermediate species and chain terminating reactions makes it an important species for modelling the low temperature oxidation of DEE. At this stage a seed mechanism is also introduced for ethanol, based on the work of Zhang et al.,¹³⁸ containing 74 reactions. This seed mechanism is included to guide the generation of ethanol specific chemistry present during DEE combustion. Ethanol is a known product of DEE decomposition at high temperatures. JSR-GC measurements of DEE oxidation performed by Tran et al.¹⁰ also show the presence of ethanol at temperatures of 500–1000 K, indicating some formation due to low temperature chemistry. This additional seed mechanism also enables the prediction of blended DEE and ethanol fuel, as shown in Section 4.4.

After mechanism generation, the model is subject to a suite of post-processing and analysis techniques, which are utilized to further improve the predictive capabilities of the mechanism. The initial post-processing step is the identification of non-physical reaction rates which violate the bimolecular collision rate limit. These collision rate violators are common in modern detailed mechanisms, as shown in the study of Chen et al.¹³⁹ Their work investigated 20 models for such violators and found that 15 of these contained either a large number of offending reaction rates or rates which exceeded the limit by a considerable factor. The number of violators also increases as the reaction rate parameters present in seed mechanisms and libraries are extrapolated to conditions outside of their validated regimes. Fortunately, tools for the screening of collision rate violators are provided by RMG, which generates a

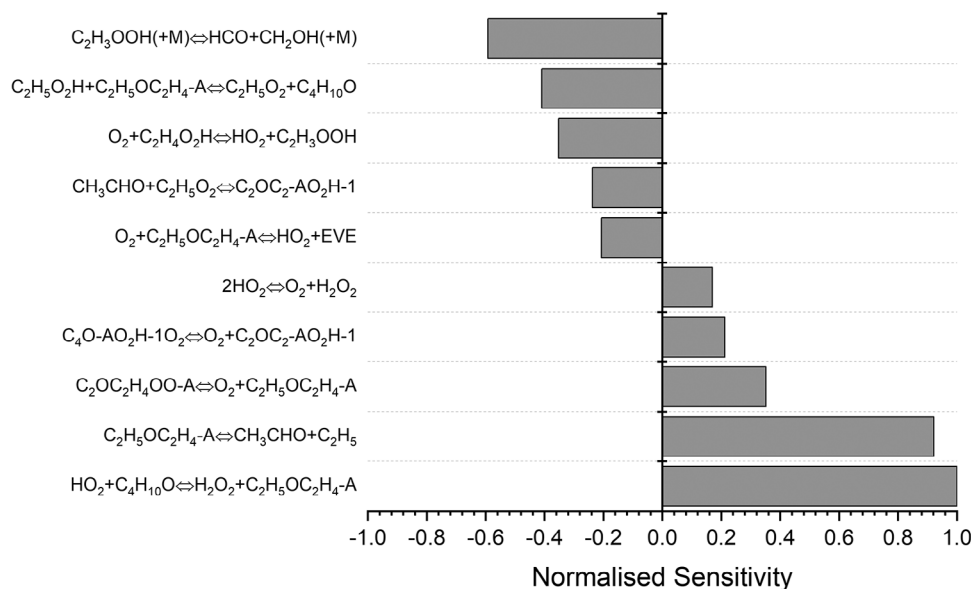


FIGURE 3 An example showing the top 10 most sensitive reactions as identified by a local A-factor sensitivity analysis for peak OH concentration predictions at $T_c = 700$ K, $P_c = 20$ bar, and $\phi = 1.0$. A positive sensitivity indicates that an increase in the reactions A-factor would produce an increase in OH concentration, which corresponds with a decrease in IDTs.

report of violating reactions after mechanism generation. However, it is the user's responsibility to act on this report. For the mechanism generated in stage (4), reactions which violate the bimolecular collision rate are removed from the seed mechanism (or library) and the mechanism is re-generated. In total, only six violators are identified for the DEE mechanism generated in this stage, as most of the potentially violating reactions are removed from seed mechanisms prior to generation, due to the exclusive selection of high-quality rate parameters. These reactions only violate the collision limit at the extremes of the investigated thermodynamic regime, at temperatures of 500 and 1500 K, and pressures of 1 and 40 bar. Of the six collision limit violators, only the decomposition of ethanol into CH_2CH_2OH and a H radical exceeded the rate limit by a considerable amount (a violation factor of 681.28) and only at conditions of 500 K and 1 bar.

As described earlier, reaction rate and thermodynamic sensitivity analysis are utilized to identify sensitive reactions and species, with sensitive and uncertain values being targeted for improvement. RoP analysis is combined with knowledge of low temperature oxidation chemistry and species or group specific literature, to identify any unusual, high flux reaction paths or the absence of expected pathways. Using this information, it is then possible to investigate the reaction seed or database to determine the source of these erroneous reactions or determine why pathways may be missing. For example, it is not uncommon for inaccurate thermochemistry or an inappropriate database reaction to make an unrealistic reaction preferential to the accepted consensus pathways. All the reaction rate parameter changes performed as a result

of these analyses are presented in [Supplementary Materials](#). An example of a local A-factor sensitivity analysis for peak OH concentration is shown in Figure 3, which displays the normalized sensitivities for top 10 most sensitive reactions at conditions of $T_c = 700$ K, $P_c = 20$ bar, and $\phi = 1.0$. In total, 39 reactions are subject to changes and eight reactions are added to the mechanism. Of particular importance for the prediction of IDTs is the removal of an RMG generated reaction, from the template family 'R_Addition_Multibond': $CH_3CHO + C_2H_5O_2 \rightleftharpoons C_2OC_2-AO_2H-1$. This reaction describes the reverse of an elementary step, which is already accounted for by a lumped QSSA reaction present in the seed mechanism, for the C-O and C-C β -scission of QOOH species ($C_2OC_2-AO_2H-1 \rightleftharpoons CH_3CHO + CH_3CHO + OH$).¹⁰ Reactions describing the 2nd oxygen addition and subsequent O₂QOOH decomposition to form ketohydroperoxides and an OH radical, are modified with the rates proposed by Danilack et al.,⁹⁰ to account for updates made to sensitive thermochemistry, identified via an enthalpy of formation sensitivity analysis for IDTs. Of the sensitive reactions displayed in Figure 3, the reaction $O_2 + C_2H_5OC_2H_4-A \rightleftharpoons HO_2 + EVE$ is also modified, with the RMG generated rate being updated using the ethanol oxidation rate parameters of da Silva et al.¹⁴⁰

The Tran et al.¹⁰ mechanism, which provided the DEE seed and thermochemistry, utilized the modified O₂QOOH thermochemistry of Sakai et al.⁴⁹ for species C₄O-AO₂H-BO₂, C₄O-AO₂H-1O₂, C₄O-AO₂H-2O₂, C₄O-BO₂H-AO₂, C₄O-BO₂H-1O₂, and C₄O-BO₂H-AO₂. Thermochemistry for these species is extracted from the Tran et al.¹⁰ mechanism and included in the DEE

TABLE 3 A catalogue of thermochemistry modifications made to the RMG generated mechanism. The literature sources for each set of species thermochemistry are provided in the table.

Species name	Structure	Original $\Delta H_{f,298K}$ (kcal/mol)	Updated $\Delta H_{f,298K}$ (kcal/mol)
EVE		-33.80 ¹⁰	-34.43 ⁹⁰
C4O-AO2H-BO2		-75.50 ⁴⁹	-71.66 ⁹⁰
C4O-AO2H-IO2		-86.04 ⁴⁹	-83.63 ⁹⁰
C4O-AO2H-2O2		-75.42 ⁴⁹	-74.13 ⁹⁰
C4O-BO2H-AO2		-75.32 ⁴⁹	-71.08 ⁹⁰
C4O-BO2H-IO2		-75.57 ⁴⁹	-72.68 ⁹⁰
C4O-BO2H-2O2		-63.70 ⁴⁹	-72.68 ⁹⁰

thermochemistry library created for this work. These species are identified as sensitive for IDT predictions (as has been discussed previously in the literature^{37,49}) and uncertain, due to the -10 kJ/mol enthalpy of formation change made by Sakai et al.⁴⁹ This change is 2.2 times larger than the mean absolute error reported for the CBSQB3 method used in their calculations (4.6 kJ/mol).^{49,141} Specific thermochemistry for these species is replaced in the current work with the calculated values of Danilack et al.,⁸⁶ which applied a higher level of theory (UCCSD(T)-F12b/cc-pVTZ-F12//B2PLYPD3/cc-pVTZ). The thermochemistry of EVE is also found to be sensitive for IDT predictions but the difference between the enthalpy of formation as determined by CBSQB3 methods in the Tran et al.¹⁰ mechanism, and the value calculated by Danilack et al.⁸⁶ is small (0.63 kcal/mol or 2.64 kJ/mol). It should be noted that neither reaction rate parameters nor species thermochemistry have been 'tuned' in this work in order to achieve accurate model predictions. Model updates are based entirely on reducing the uncertainty of sensitive thermokinetic parameters by updating these with recent literature data of a lower degree of uncertainty. Any predictive improvements are a by-product of this methodology. Thermochemistry modifications are summarized in Table 3.

After mechanism generation and post-processing modification, the model is reduced using the Ansys Workbench software and the direct relation graph with error propagation (DRGEP) reduction method.¹⁴² This reduction is performed using a zero-dimensional homogeneous reactor to simulate IDTs, at conditions of $\phi = 0.5, 1.0,$ and $2.0, P_c = 1, 10, 20,$ and 40 bar, and $T_c = 500-1100$ K. The reduction is performed with an IDT tolerance of 3% throughout the temperature regime, which translates to a flux tolerance of 2.25%, to remove any superfluous reactions included by the RMG generation. Due to the flux criteria employed during generation, it is common for RMG to generate mechanisms which contain a large fraction of such reactions, which are not important to the overall combustion process. In this case, the mechanism reduction process produces a mechanism which contains 146 species and 4392 reactions (a reduction of 95 species and 3037 reactions from the full mechanism). Of these reactions, rate parameters for 140 reactions are included directly from seed mechanisms (86 reactions from the DEE seed mechanism, 72 from the ethanol seed mechanism, and 40 from the small molecule core), 400 reactions are sourced from the Aramco 2.0 reaction library, eight additional reactions are included during post-processing, and the remaining 3844 reactions are generated by RMG (1050

are generated as pressure dependent networks and 2794 are estimated based RMG reaction families and training reactions). Clearly, most of the reactions present in the RMG mechanism are generated using training reaction rate parameters (>62% of the reactions in the reduced mechanism). It would not be feasible to calculate or determine reaction rate parameters for each of these reactions using experimental or quantum mechanical methods when attempting to produce a mechanism. For comparison, the mechanism of Tran et al.¹⁰ contains 746 species and 3555 reactions, while that of Sakai et al.⁴⁹ contains 341 species and 1867 reactions, and that of Danilack et al.⁹⁰ contains 228 species and 1377 reactions. Hence the reduced RMG mechanism contains the fewest species but still has the largest number of reactions. This may partly be due to some remaining low importance reactions, but it is also likely that 'by-hand' mechanisms will contain lumped reaction steps, whereas RMG produces elementary reaction sequences which leads to a larger number of overall steps. For simulation purposes, it is the number of species which is more critical to the computational cost of solving the chemical rate equations numerically.

As shown in Figure 2, the application of a combination of the methods applied in stages (2) and (3), alongside post-processing and analysis, produces a mechanism (development stage (4)) which provides extremely accurate IDT predictions throughout the entire temperature regime. The NTC region displays the correct intensity and occurs at the appropriate temperature. It also features the 'double hump' or 'shouldering' which is indicative of the multi-stage ignition typical of DEE autoignition,⁶⁷ and is reliant on various key chain branching O-O bond scission reactions, as shown in the recent work of Sakai et al.¹⁴³ However, it is not sufficient for a mechanism to be validated by a global target (such as IDTs) over a small range of conditions ($\varphi = 1.0$, $P_c = 20$ bar, $T_c = 500$ – 1100 K) if it is to be generalizable and applicable for simulations of different types of combustors. A good, detailed mechanism must be robust, providing accurate predictions for a range of targets at many engine relevant conditions.

4 | MODEL EVALUATION

In the following sections, the RMG generated mechanism produced in development stage (4) is evaluated against various experimental DEE target measurements (as detailed in Section 2.2). However, the objective of this study is not primarily to produce an effective and robust DEE model. As previously discussed, the primary objective of this study is to investigate to capabilities of AMG (specifically RMG), for the production of detailed advanced biofuel mechanisms, using DEE as a case study

fuel species. Therefore, this evaluation will not focus on a detailed analysis of the underlying chemistry driving the autoignition of DEE, which has already been thoroughly documented in the literature.^{10,67,72,86,90,144} Instead, the following sections provide a critical analysis of the RMG models predictive capabilities, relative to a selection of literature sourced mechanisms.^{10,49,90} These literature mechanisms are selected as, while they may borrow some DEE specific reactions and thermochemistry from previous mechanisms (including each other), they have been independently produced and include a variety of different reaction pathways, rate parameters, and thermochemistry. They have also been produced for different purposes and as such, are validated against different target parameters. The Sakai et al.⁴⁹ mechanism was evaluated against RCM and ST IDTs by the original authors, whereas Tran et al.¹⁰ evaluated their mechanism against speciation measurements taken from JSR and PFR experiments. The Danilack et al.⁹⁰ mechanism has not been validated against any experimental results, as the objective of their work was not to provide highly accurate DEE predictions, but to quantify the effects of non-Boltzmann reactions on predictions as well as the impact of often ignored diastereomers.^{86,90} This mechanism is included in the comparative assessment due to the large amount of newly generated thermochemistry and reactions rate coefficients produced, which were calculated at a relatively accurate level of theory (UCCSD(T)-F12b/cc-pVTZ-F12//B2PLYPD3/cc-pVTZ) some of which have been incorporated into the final mechanism produced in this work.

4.1 | Diethyl ether ignition delay times

The homogeneous IDT of a fuel is not only important for determining knock resistance and pre-ignition propensity. It is a highly important property which defines the time taken for a specific fuel mixture to oxidize and engage in rapid heat release, at a given set of thermodynamic conditions. The duration of this delay greatly influences the evolution of the combustion process within reciprocating engines.^{9,145} Therefore, it is a bare minimum requirement that an automatically generated detailed model, designed for the prediction of engine relevant combustion phenomena, can accurately predict the associated IDTs. Figure 4 shows the variable volume model predictions of the RMG mechanism generated in this study, compared alongside the predictions of three literature mechanisms^{10,49,90} and the RCM and ST measurements of Issayev et al.⁶⁷ and Uygun.⁷⁰ The conditions presented range from $\varphi = 0.5$ – 1.0 , $P_c = 20$ – 40 bar, and $T_c = 550$ – 1100 K.

At both stoichiometric conditions ($P_c = 20$ and 40 bar), the IDT predictions of the RMG generated mechanism are

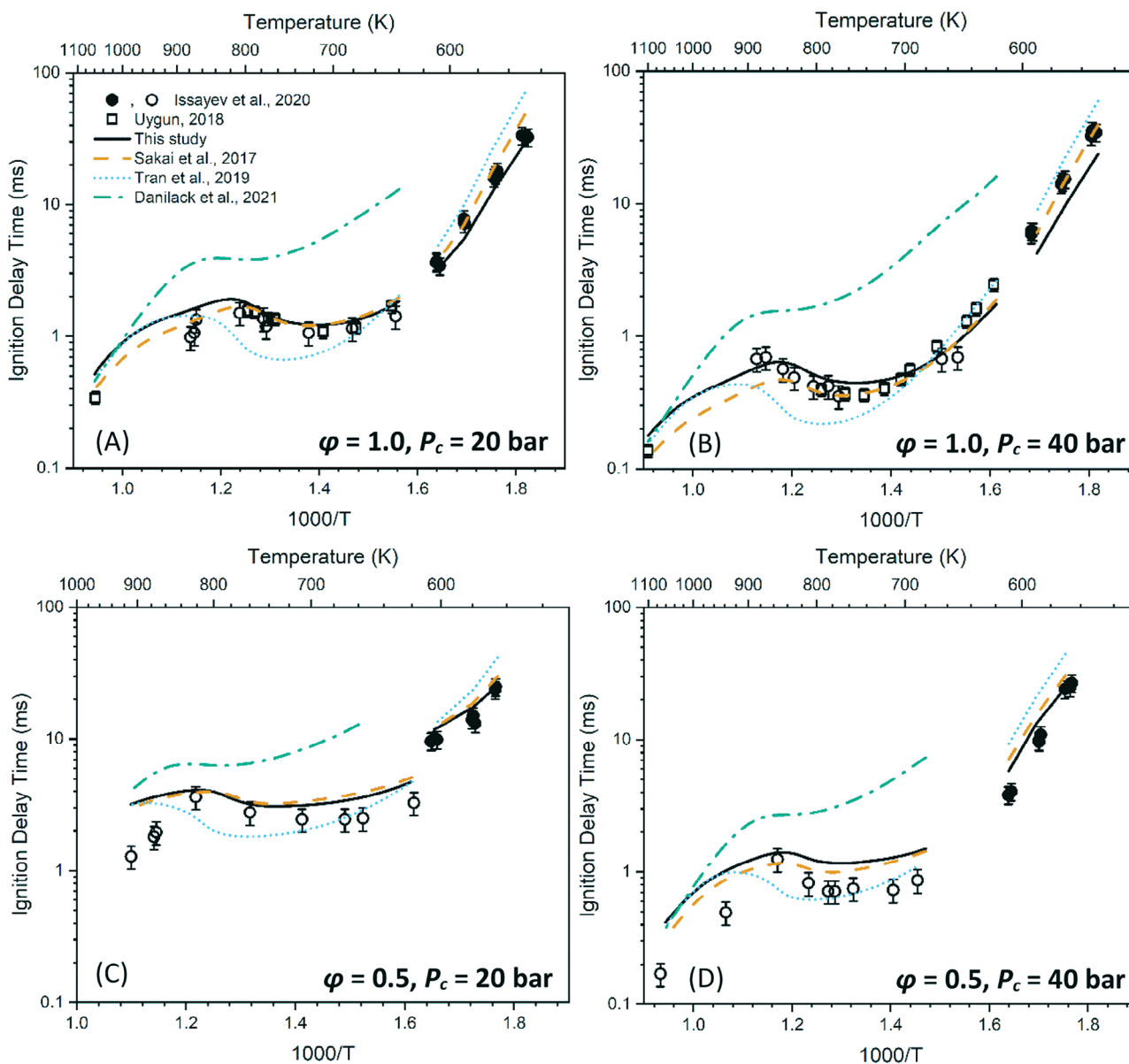


FIGURE 4 A comparison of predicted IDTs from this study and selected literature sourced mechanisms,^{10,49,90} accounting for RCM and ST facility effects. Predictions are evaluated against the experimental data of Issayev et al.⁶⁷ and Uygun.⁷⁰ Closed symbols represent RCM measurements and open symbols show ST measurements.

excellent. At 20 bar, the RMG model produces IDTs which are within the uncertainty range of experimental measurements throughout the temperature regime, with the exception of the highest temperature condition. However, here the difference between experimental measurements and simulations is still minor (<0.2 ms). As previously mentioned, the predictions in the NTC region are particularly impressive. The story is similar for the 40 bar case, with only a small over prediction of reactivity for the low temperature RCM predictions. When comparing against the predictions of the selected literature mechanisms, only the Sakai et al.⁴⁹ mechanism provides a similar level of accuracy, particularly in the NTC region. While the IDT

predictions of the Sakai et al.⁴⁹ mechanism are marginally more accurate at 20 bar, it fails to match the NTC intensity observed at 40 bar. The RMG mechanism matches this intensity accurately. The Tran et al.¹⁰ mechanism provides reasonable predictions at low and high temperatures for both stoichiometric cases, but fails to match the observed NTC behavior for either pressure. Notably, the RMG mechanism out-performs the Tran et al.¹⁰ mechanism for these IDT targets, although it provided the DEE seed during mechanism generation. As expected, the mechanism of Danilack et al.⁹⁰ does not provide accurate IDT simulations, under-predicting reactivity throughout. Variable volume RCM predictions are not shown for the Danilack

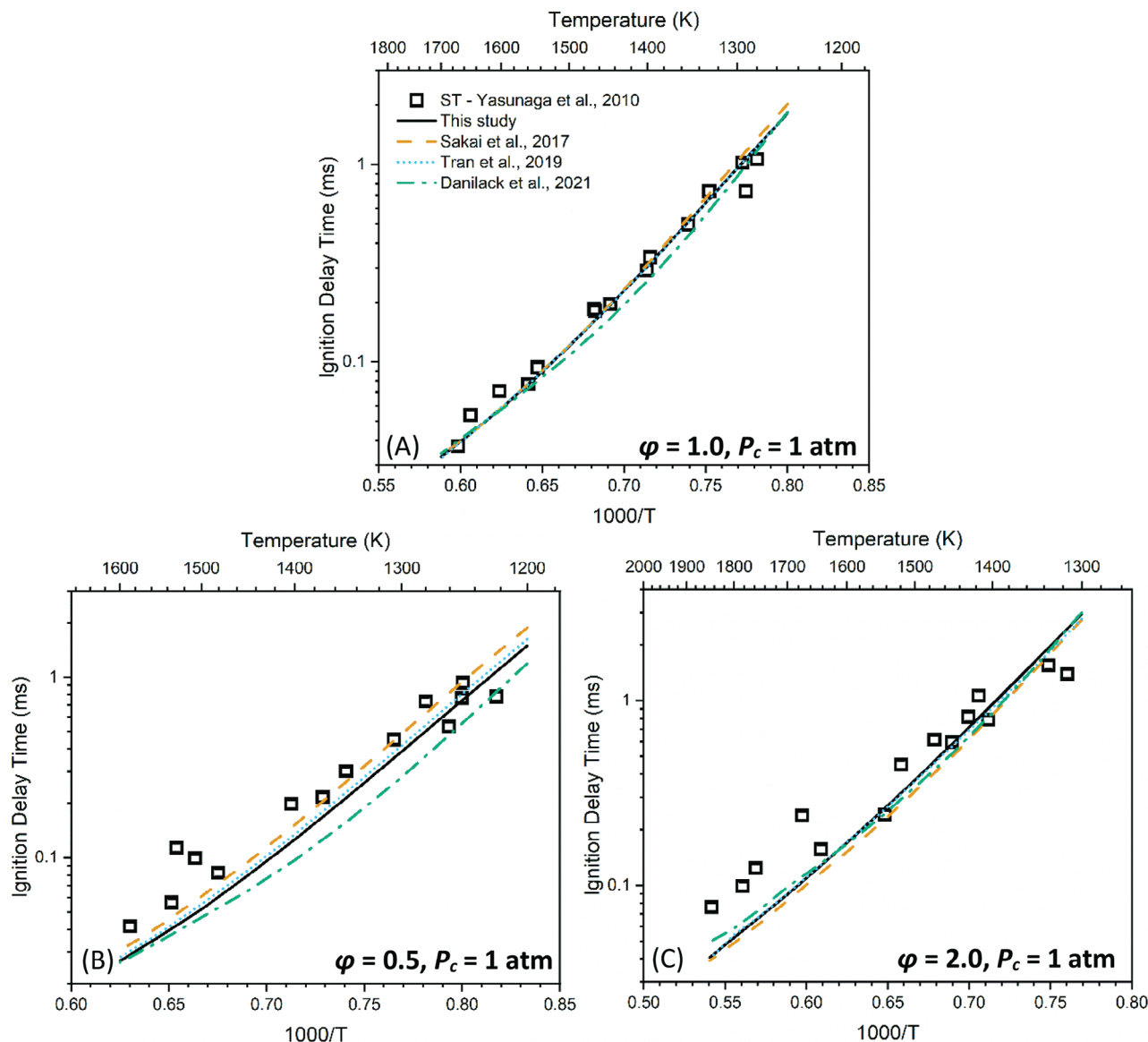


FIGURE 5 High temperature DEE IDTs, as predicted by the RMG model and selected literature mechanisms. Predictions are compared against the ST measurements of Yasunaga et al.,⁷² shown by open symbols. All fuel mixtures constitute 1% DEE by mole in an argon diluent, with varying O₂ content, dependent on stoichiometry.

et al.⁹⁰ mechanism at any of the conditions displayed in Figure 4, as all of these predicted IDTs in excess of 100 ms.

Under lean ($\phi = 0.5$) fuel/air conditions, the models all lose some accuracy at the highest temperatures. The IDT predictions of the RMG and Sakai et al.⁴⁹ models are nearly identical throughout both the 20 and 40 bar lean regimes, providing good predictions of the low temperature RCM measurements and reasonable predictions for NTC region ST measurements. The mechanism of Tran et al.¹⁰ however, tends to under-predict DEE reactivity for RCM measurements and over-predict reactivity in the NTC region. Overall, both the RMG model and the Sakai et al.⁴⁹ mechanism display good predictive capabilities for a wide range of engine relevant conditions, with minor perfor-

mance differences between the two models. However, it should be noted that the RMG mechanism contains less than half the amount of species contained in the Sakai et al.⁴⁹ mechanism (RMG: 146, Sakai et al.⁴⁹: 341).

At higher temperatures (>1200 K), the IDT predictions of the models are all similar. This can be seen in Figure 5, which compares the predictions of the models with the ST measurements of Yasunaga et al.,⁷² at 1 atm and $\phi = 0.5$, 1.0, and 2.0. The similarity between the predictions of the RMG, Sakai et al.,⁴⁹ and Tran et al.¹⁰ mechanisms is likely due to the similar small molecule chemistry, which such high temperature IDTs are known to be highly dependent on. The Danilack et al.⁹⁰ mechanism shows some minor differences when compared to the others, which

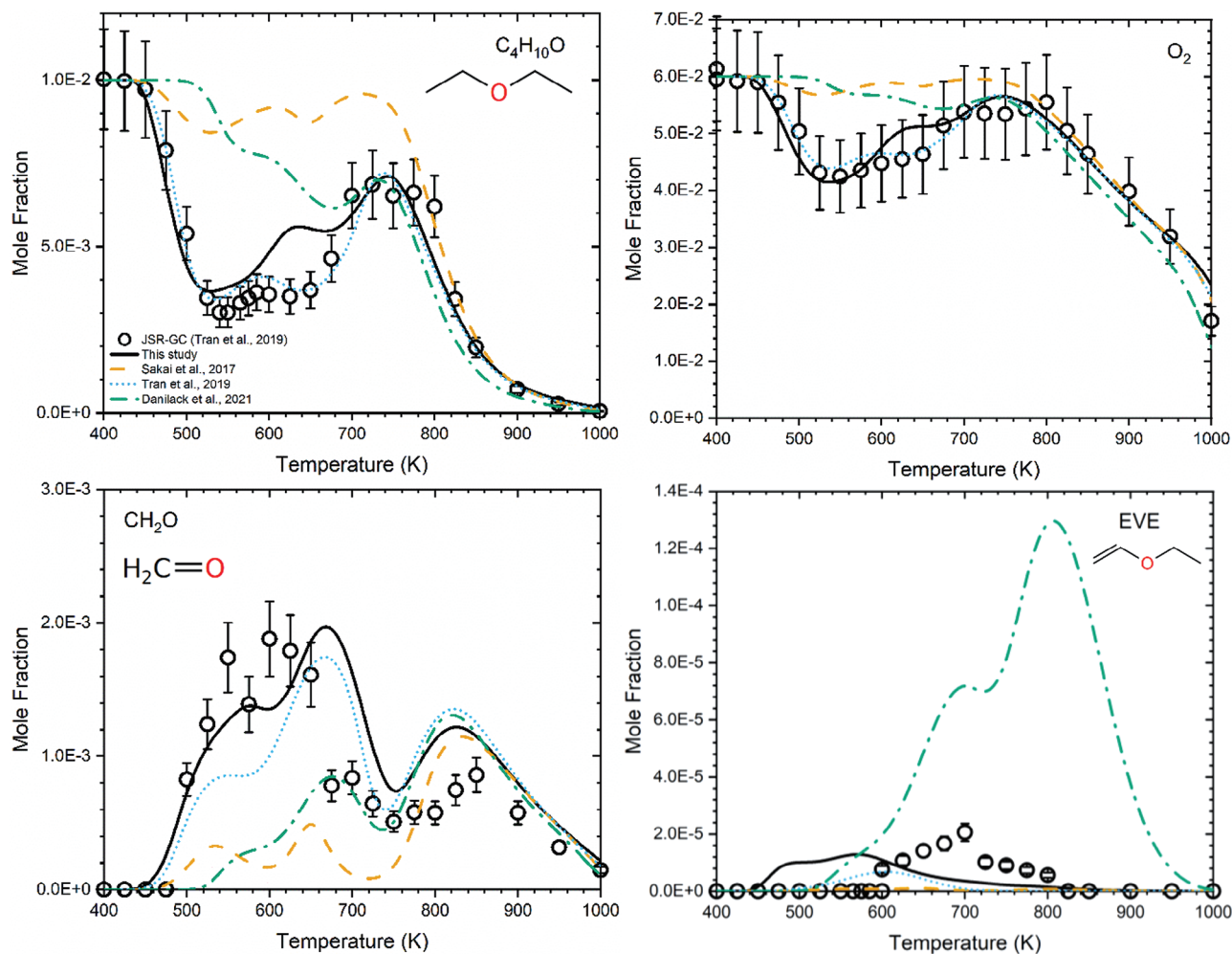


FIGURE 6 Model predictions for the concentrations of DEE ($C_4H_{10}O$), O_2 , CH_2O , and EVE, in a JSR with varying temperature. Symbols show the JSR-GC measurements of Tran et al.¹⁰ Stoichiometric, 106.7 kPa, 2 s residence time, 1/6/93 mol% DEE/ O_2 /He.

are possibly a result of the new small molecule chemistry employed.^{86,146} In general, these high temperature ST IDT predictions are reasonable for all the investigated models. However, in a similar trend to that observed in Figure 4, stoichiometric predictions are clearly the most accurate. Changing the stoichiometry to either lean or rich decreases the effectiveness of all the models, indicating that more fundamental research is required at these conditions to determine accurate reaction rate parameters.

4.2 | Jet stirred reactor species concentrations

Species concentration measurements obtained during fuel oxidation experiments can provide valuable insight into a fuel's combustion chemistry, as well as identifying potentially toxic or harmful combustion products which may limit the viability of an advanced biofuel. It is therefore important that such species concentration measurements

can be well predicted by a model if it is to inform engine and fuel blend design. Providing only an assessment of IDT predictions would also constitute a misrepresentation of the RMG model's capabilities relative to the literature sourced mechanisms, particularly as the Tran et al.¹⁰ model was primarily developed to model JSR-GC measurements, rather than IDTs. Figure 6 shows the JSR-GC measurements of Tran et al.,¹⁰ for DEE, O_2 , CH_2O , and EVE. The predictions of the four models are also shown.

DEE concentration predictions are very good for the RMG generated mechanism and display the characteristic double NTC. However, there is a clear under-prediction for DEE conversion from 600 to 700 K. This is in contrast to the Tran et al.¹⁰ mechanism which predicts each JSR-GC measurement almost exactly, including this low temperature NTC region. Even with this difference, the RMG predictions are still good, especially considering that the mechanism contains less than 20% of the number of species present in the Tran et al.¹⁰ mechanism. The other literature sourced mechanisms do not provide good

predictions for the concentrations of DEE over the temperature range but do display similar double NTC behavior. The study of Tran et al.¹⁰ identified that the first NTC results from competition between the β -scission reactions of hydroperoxyl (QOOH) fuel radicals and the first oxygen addition. During the development of the RMG mechanism, the first oxygen addition was identified as sensitive for IDT prediction. The Tran et al.¹⁰ rates for these reactions are reverse rates simply calculated using the Sakai et al.⁴⁹ rates and thermochemistry, originally calculated at the CBSQB3 level of theory. To eliminate the potential for uncertainty introduced by this reverse rate calculation, these rates were reverted to the values presented by Sakai et al.⁴⁹ This modification contributed considerably to both the under-prediction of DEE consumption during this NTC but also the accurate prediction of NTC IDTs. Clearly, more fundamental work is required here to determine accurate rate parameters and thermochemistry, beyond composite CBSQB3 calculations.

The predictions of the RMG mechanism are good for O₂ concentrations and are very similar to the predictions of the Tran et al.¹⁰ mechanism. Formaldehyde (CH₂O) is also important not only due to its environmental consequences on emission, but it is also thought to contribute significantly to heat release behavior during oxidation, as it is present in large quantities and is quickly consumed as a precursor to HO₂ formation.^{147,148} At these conditions, the RMG mechanism produces the most accurate predictions for CH₂O concentration. At temperatures of 500–600 K, the Tran et al.¹⁰ considerably underpredicts the formation of CH₂O, whereas the RMG mechanism gives predictions much closer to the measured values. When temperatures exceed 800 K, all of the investigated models over-predict the formation of CH₂O, but the RMG model provides a marginally better estimate than the Tran et al.¹⁰ mechanism. While the predictions of the Sakai et al.⁴⁹ mechanism are marginally closer to the measured value at this temperature than the RMG mechanism, the predictions throughout the rest of the temperature regime are poor. None of the mechanisms produce accurate predictions for EVE concentrations. For the RMG mechanism, most of the reactions associated with EVE formation (such as the hydrogen abstractions of fuel radical species) and consumption are generated by RMG rate rules, introducing a high degree of uncertainty into these predictions. However, the predictions of the RMG model are the closest to the experimental values in terms of magnitude, suggesting that the database expansion and inclusion of an EVE sub-mechanism has been beneficial. Dewey and Rotavera¹⁴⁹ discussed the importance of the consumption mechanisms for R+O₂ intermediates and highlighted the benefits of including sub-mechanisms for such species in the RMG generation process. Their study found that the

inclusion of these sub-mechanisms led to an improvement in model predictions of experimental species profiles, similar to the results shown in this study for the RMG generated mechanism. Rate parameters for initiation through the hydrogen abstraction of DEE by an OH radical are not estimated by RMG and are identical to those provided in the Tran et al.¹⁰ seed mechanism. Hartness and Rotavera¹⁵⁰ recently identified the significance of branching fractions for such initiation reactions in mechanisms which include peroxy radical chemistry. The RMG generated mechanism used in this study and the mechanism of Tran et al.¹⁰ provide the best estimates for EVE concentration, while sharing the same rate parameters for the highly influential hydrogen abstraction by OH reactions. However, it is difficult to form conclusions based only on these reactions, due to the considerable differences between each mechanism investigated, including significant differences in the rate parameters and thermochemistry of several influential reactions and species, respectively. Total reaction rate coefficients and branching ratios for these reactions are provided in [Supplementary Materials](#) for each of the literature sourced mechanisms. The formation and consumption of EVE is clearly a target for future DEE model development in general, not just for the RMG generated mechanism, as this species is associated with chain terminating low temperature oxidation pathways for DEE.

Many species concentrations show a bimodal distribution with changing temperature, including C₂H₄, C₂H₅OH, and C₄H₈O₂-CY, as shown in Figure 7. The RMG mechanism successfully predicts such a distribution for the concentration of C₂H₄ but, like the Tran et al. mechanism,¹⁰ the first concentration peak is considerably over-predicted. The Sakai et al.⁴⁹ and Danilack et al.⁹⁰ mechanisms give a good prediction of C₂H₄ concentrations in this initial peak and all the investigated mechanism predict the higher temperature second peak accurately. C₂H₅OH is formed in small quantities during the low temperature oxidation of DEE. Below 600 K the RMG model significantly under-predicts the concentrations of C₂H₅OH, whereas the Tran et al.¹⁰ mechanism predicts concentrations well in this region. However, this is not the case at temperatures >600 K, wherein only the RMG mechanism provides reasonable predictions. C₄H₈O₂-CY is a cyclic ether formed by the decomposition of QOOH species to produce C₄H₈O₂-CY and an OH radical. Again, the RMG mechanism produces the most accurate predictions for the concentration of this species, with the Tran et al. mechanism under-predicting the concentration between 575 and 675 K. The bimodal concentration distribution is predicted by both the RMG and Sakai et al.⁴⁹ mechanisms, although in both cases the location of the second concentration peak occurs at a lower temperature than

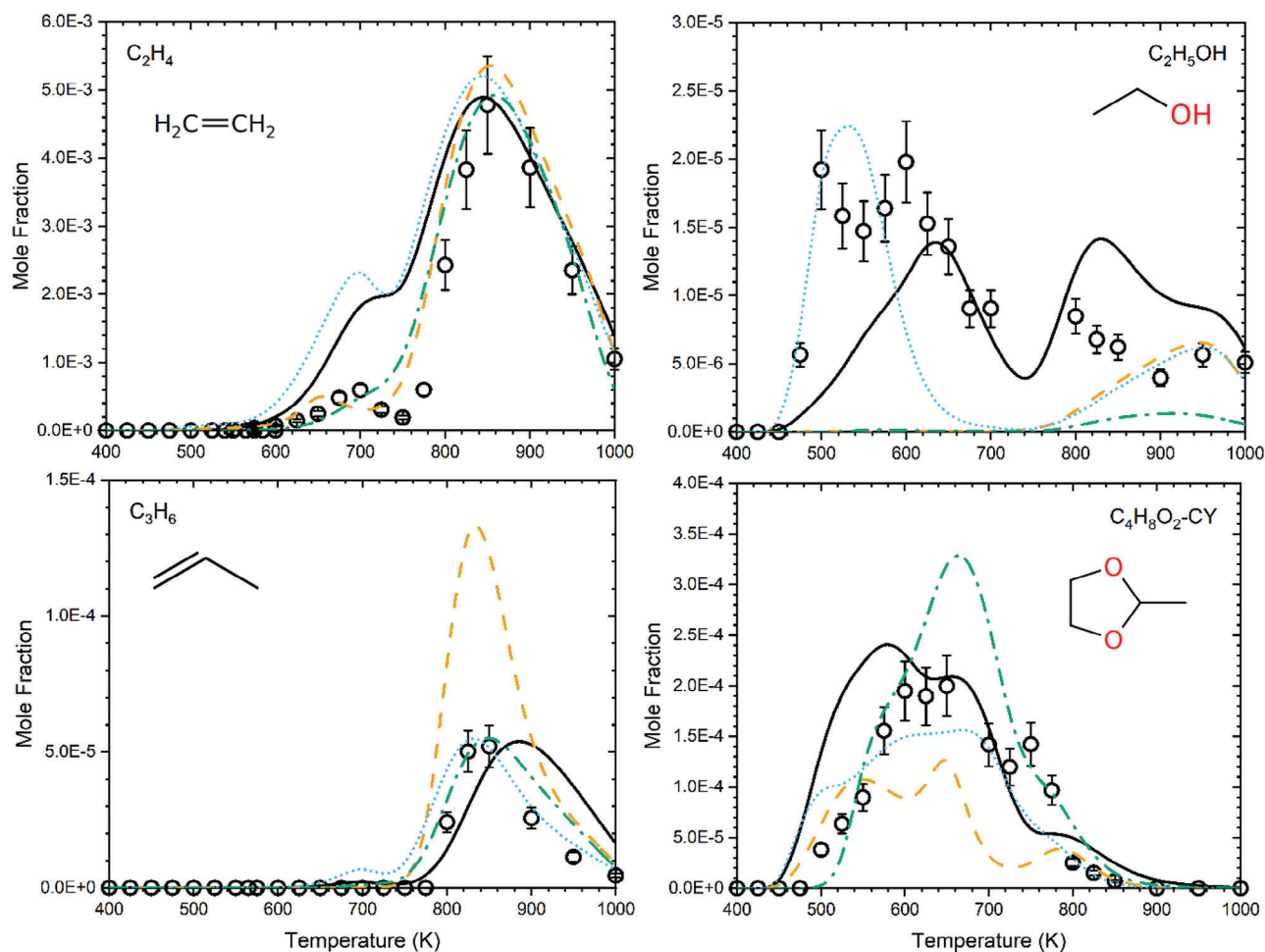


FIGURE 7 Model predictions for the concentrations of ethene (C_2H_4), ethanol (C_2H_5OH), propene (C_3H_6), and 2-Methyl-1,3-dioxolane ($C_4H_8O_2$ -CY), in a JSR with varying temperature. Symbols show the JSR-GC measurements of Tran et al.¹⁰ Stoichiometric, 106.7 kPa, 2 s residence time, 1/6/93 mol% DEE/ O_2 /He.

is observed experimentally. While the Tran et al.¹⁰ mechanism predicts also predicts a bimodal distribution for C_3H_6 concentrations, this is not observed in the experimental data, nor is it predicted by any of the other models.

As pressures are increased to more engine relevant conditions, the species concentration predictions of the RMG mechanism continue to show a good level of accuracy, comparable to those of the Tran et al.¹⁰ mechanism. Species concentrations for DEE and CH_2O at lean fuel/air conditions ($\phi = 0.5$) and 10 atm are displayed in Figure 8. It is clear to see that the increase in pressure results in a suppression of the DEE NTCs, though some NTC behavior is still apparent. It is suggested in the literature that the suppression of the first NTC is due mainly to a promotion of 2nd oxygen addition reactions and the high-pressure stabilization of RO_2 and $QOOH$ radicals. The 2nd NTC is also impacted by high-pressure RO_2 stabilization but is also suppressed pressure dependent enhancement of HO_2 chemistry.¹¹⁰ The RMG mechanism correctly predicts this suppression but still marginally over-predicts DEE concen-

trations in both NTC regions. However, the predictions of the RMG mechanism are still much more accurate than those of the Sakai et al.⁴⁹ and Danilack et al.⁹⁰ mechanisms and provide a reasonable degree of accuracy throughout the temperature regime. CH_2O concentration predictions are good for both the RMG mechanism and the Tran et al.¹⁰ mechanism, with the RMG simulations more accurately predicting concentrations during the 700 K peak and the Tran et al.¹⁰ mechanism achieving a marginally better accuracy in the 850 K region.

When the pressure is increased further to 100 atm (a pressure not uncommon during combustion within reciprocating engines), the suppression of DEE NTC behavior is even more apparent. At the lean ($\phi = 0.5$), high-pressure conditions displayed in Figure 9, the experimental measurements of Wang et al.¹¹⁰ show none of the characteristic low temperature combustion behavior. The RMG model predicts this change in behavior well, with only the Tran et al.¹⁰ mechanism providing marginally more accurate predictions throughout the temperature regime.

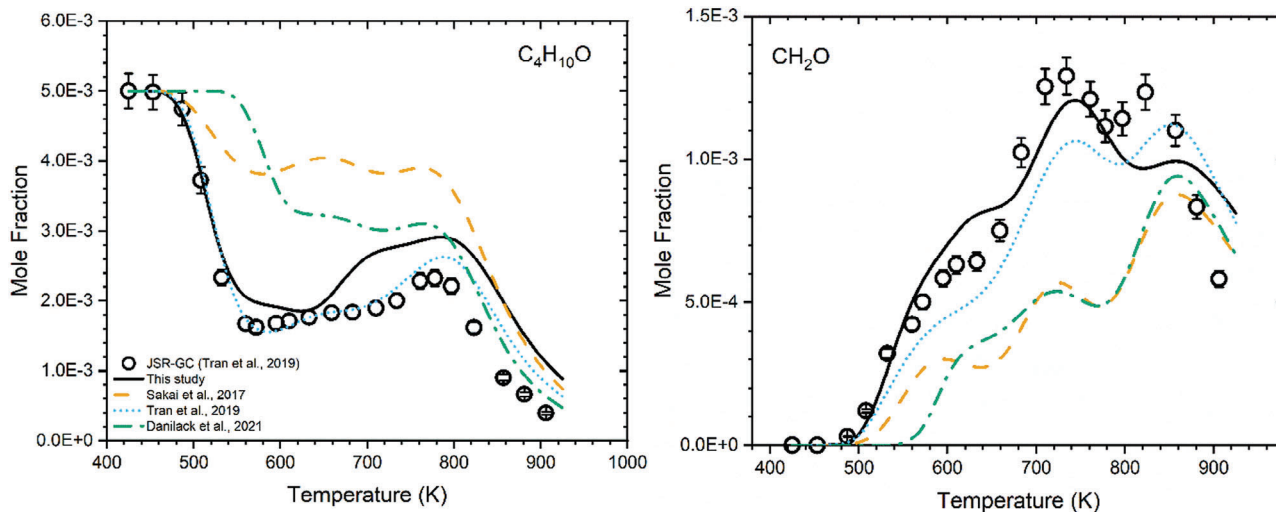


FIGURE 8 Model predictions for the concentrations of DEE ($C_4H_{10}O$) and CH_2O , in a JSR with varying temperature. Symbols show the JSR-GC measurements of Wang et al.¹¹⁰ Lean ($\phi = 0.5$), 10 atm, 1.2 L/min flow rate, 0.5/6/93.5 mol% DEE/ O_2/N_2 .

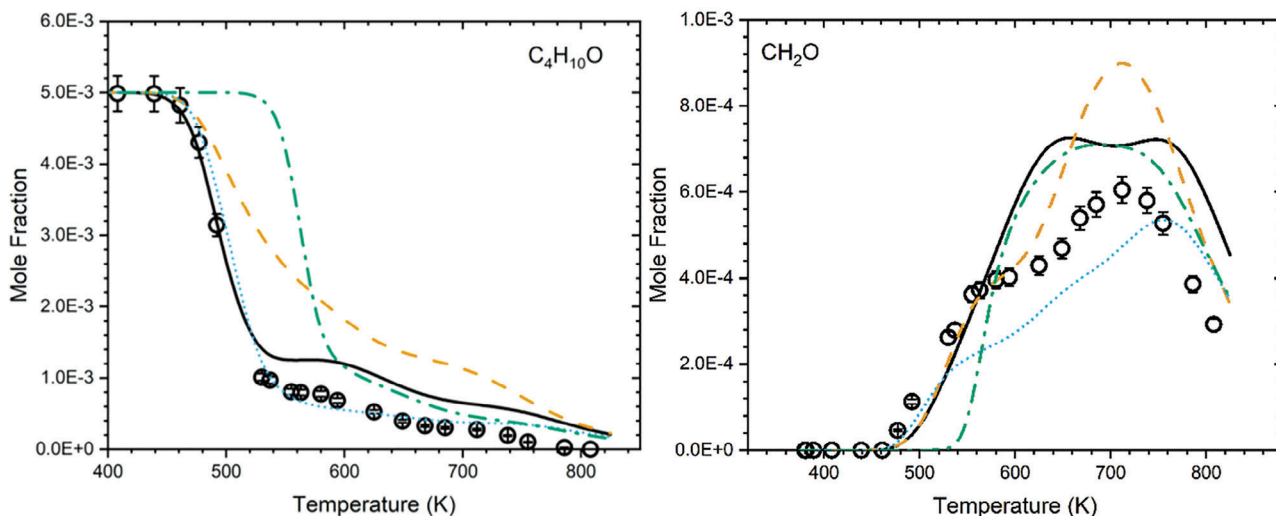


FIGURE 9 Model predictions for the concentrations of DEE ($C_4H_{10}O$) and CH_2O , in a JSR with varying temperature. Symbols show the JSR-GC measurements of Wang et al.¹¹⁰ Lean ($\phi = 0.5$), 100 atm, 3 L/min flow rate, 0.5/6/93.5 mol% DEE/ O_2/N_2 .

CH_2O predictions are reasonable for all the investigated mechanisms, though the Sakai et al.⁴⁹ mechanism does over-predict CH_2O concentrations considerably around 700 K. The RMG and Danilack et al.⁹⁰ mechanisms also over-predict CH_2O concentrations after the initial peak at 550 K, whereas the Tran et al.¹⁰ mechanism under-predicts concentrations for the majority of the temperatures investigated.

4.3 | Diethyl ether flame speeds

Adiabatic DEE flame speeds, as measured by Gillespie et al.,⁸⁰ are displayed in Figure 10 alongside the

predictions of the RMG, Tran et al.,¹⁰ and Danilack et al.⁹⁰ mechanisms. The Sakai et al.⁴⁹ mechanism is not included in this assessment, as no transport data are provided by the authors, and to generate new transport data would be misrepresentative of the mechanism's performance. As supplied, the Danilack et al.⁹⁰ mechanism also contains incomplete transport data, missing parameters for eight DEE oxidation intermediates: the two ROOH species ($C_2OC_2H_4OOH-A$ and $C_2OC_2H_4OOH-B$), ethoxyacetaldehyde ($C_2H_5OCH_2CHO$), both possible ethoxyethanols ($C_2H_5OC_2H_2OH-A$ and $C_2H_5OC_2H_4OH-B$), ethyl acetate (EA), and both RO species ($C_2H_5OC_2H_4O-A$ and $C_2H_5OC_2H_4O-B$). These species are only involved in reactions which Danilack et al.⁹⁰ sourced from the Tran

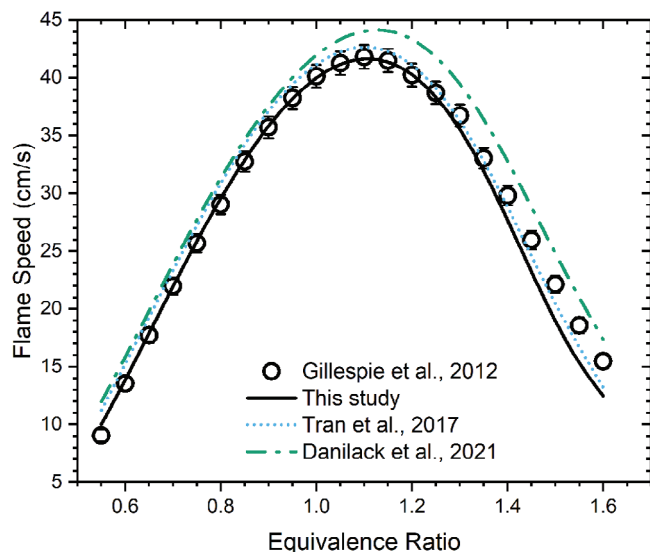


FIGURE 10 A comparison of the 298 K DEE flame speed measurements of Gillespie et al.,⁸⁰ with the model predictions of the RMG generated, Tran et al.,¹⁰ and Danilack et al.⁹⁰ Measurements and predictions are displayed for varying equivalence ratios, from $\phi = 0.55$ to 1.6.

et al.¹⁰ mechanism, so for the purposes of this analysis, the missing transport data are completed using the parameters supplied by Tran et al.¹⁰

The flame speed predictions of the RMG model are extremely accurate from equivalence ratios of $\phi = 0.55$ to 1.35. At richer equivalence ratios, the model begins to under-predict the flame speed slightly, getting worse as the fuel/air ratio increases. A similar trend is observed for the Tran et al.¹⁰ mechanism, though the predictions for both models are good throughout. While the Danilack et al.⁹⁰ mechanism produces similar predictions for lean conditions, there is a clear over-prediction of flame speeds as the equivalence ratio increases.

4.4 | Diethyl ether/ethanol blend

So far, this study has demonstrated the RMG can produce highly effective models for the prediction of DEE combustion phenomena. However, advanced biofuels are often blended with other biofuel components, fossil fuels, or combustion enhancing additives to improve engine performance or minimize harmful emissions.^{8,67,151,152} The production of a combined mechanism for such blends by-hand can be difficult and error prone. Species notation is often different between mechanisms, particularly as the intermediate oxidation species become more complex and niche, and mechanisms are not always well documented or adequately referenced. RMG generated mechanisms can be combined relatively simply by either performing

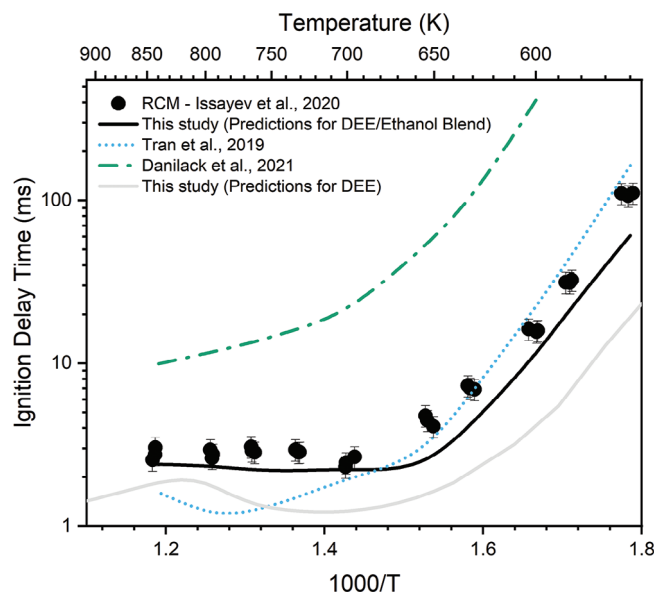


FIGURE 11 Variable volume predictions of DEE/ethanol blend IDTs, compared to the RCM measurements of Issayev et al.⁶⁷ Blending ratio of 50/50 by mole DEE/ethanol, $\phi = 1.0$, $P_c = 20$ bar, $T_c = 550$ –1100 K. IDT predictions produced using the mechanism developed in this study are also shown for DEE at the same conditions, to show the impact of DEE and ethanol fuel blending.

a direct combination (which simply combines the mechanisms together with user defined priorities, eliminating any unmarked duplicates), or by converting the individual mechanisms into seed mechanisms and using these to generate a new combined mechanism. The latter method may take considerably more computation time, but it will also generate cross reactions between the different species and their products, which a direct combination will not. As described in Section 3.4, the mechanism generated in this study utilizes the latter method described here, by including an ethanol seed mechanism based on the work of Zhang et al.¹³⁸

DEE and ethanol blends have been investigated extensively in the literature due to their strikingly different autoignition responses and common production process.^{67,153–155} The difference in reactivity between the two fuels facilitates the production of binary blends with a large range of potential octane/cetane numbers and thus may facilitate tailored fuel blending approaches. Model predictions for the IDTs of a 50/50 by mole DEE/ethanol blend are shown in Figure 11, alongside the RCM measurements of Issayev et al.⁶⁷ It is not expected that the literature models selected will provide highly accurate predictions for this blend, as they were all exclusively designed for DEE (though some ethanol specific chemistry will be covered by each mechanism's small molecule core). Simulation results for the mechanism of Sakai et al.⁴⁹ are not presented here, as convergence issues at these

conditions outside of the mechanisms validated regime prevent the completion of variable volume simulations. Predictions for the RMG generated mechanism are good throughout, indicating an accurate interpretation of the blending response between the two fuel species. The suppression of the DEE NTC by ethanol is particularly well predicted and is evident in the comparison with predictions of pure DEE IDTs in Figure 11. This is an important blending interaction which could be highly influential for determining engine performance. The high level of performance demonstrated relative to the literature mechanisms is likely due to the inclusion of an additional, high-quality ethanol seed mechanism. The literature sourced mechanisms rely on the corresponding small molecule sub-mechanisms to interpret this blending response.

5 | DISCUSSION AND CONCLUSIONS

AMG tools (such as RMG) have the potential to accelerate the production of detailed mechanisms for the combustion of complex fuels and enable the rapid design and development of advanced biofuels and strategies for their blending. Such tools use relatively simple algorithmic approaches to enable the generalization of high-quality thermochemical data through the use of reaction families, training data and group additivity methods, rather than complex machine learning tools. However, since they embed knowledge of molecular structures, they are unlikely to lead to wildly unrealistic estimates for reaction rate parameters and important thermodynamic and transport properties. Using DEE as a case study, this work has shown that when appropriate reaction families and training data are incorporated, such tools are capable of producing robust mechanisms that are comparable to, or outperform, hand-built literature mechanisms for oxygenated fuels, across a range of key combustion targets and conditions. IDT predictions for the RMG mechanism were considerably more accurate than those of the Tran et al.¹⁰ mechanism, particularly for stoichiometric conditions within the NTC region. The mechanism of Sakai et al.⁴⁹ also performed well for IDT predictions but did not predict the NTC intensity of the stoichiometric, 40 bar conditions, which was well predicted by the RMG mechanism. JSR-GC measurements were well predicted by the RMG mechanism, providing comparable results to those of the highly accurate Tran et al.¹⁰ mechanism for these targets. In the case of many intermediate oxidation species, the RMG model even outperformed the predictions of the Tran et al.¹⁰ model, as shown in Figure 6 (CH₂O and EVE) and Figure 7 (C₂H₅OH and C₄H₈O₂-CY). Speciation predictions continued to perform well as the RMG model was extrapolated beyond the conditions which it was originally designed for, as shown

by the 100 atm predictions in Figure 9. The RMG mechanism was also shown to provide accurate IDT predictions for a DEE/ethanol blend as displayed in Figure 11. No tuning of reaction rates or thermochemistry was carried out in this work.

In an ideal world, it would be possible to apply such tools with minimal user intervention from the start, in order to generate robust mechanisms. However, it was necessary in this work to update the thermochemistry database as well as training data for key reaction families relevant to complex and oxygenated functional groups. A high-quality seed mechanism was also required. However, once database updates were complete, the AMG methodology was capable of producing a model much faster when compared to the generation of such a mechanism by hand, and with considerably less input from the user. In addition, these updates will be relevant for future application of RMG to other oxygenated fuel components that are likely to become more important with the emergence of complex blended fuels, such as those produced by the alcoholysis of lignocellulosic biomass.⁸ Such blends contain multiple oxygenated functional groups (alcohol, ether, ester, and ketone groups are present in the fuel species alone) and the more complex species may have little to no data available in the literature. The database updates prepared as part of this work have been made available in [supplementary material](#) for the benefit of future applications of RMG to oxygenated fuel components. Current methods for database expansion are not always clear, making them inaccessible to inexperienced users, increasing the knowledge and time investment required to construct robust mechanisms with AMG tools. To facilitate wider application of AMG tools, and to improve the efficiency of the approach, it would be highly beneficial to have an easily accessible central database of such inputs with regular updates. Currently, such information must be sourced from the [supplementary materials](#) of published studies or GitHub branches and forks of the RMG database, which are not always adequately referenced or described. Maintaining such a database would not be without effort, but the task could potentially be automated with the help of AI tools in the future.

In summary, this study has shown that AMG tools are currently capable of producing advanced biofuel combustion mechanisms, which are robust and highly accurate across a wide range of combustion relevant targets. While the process is not currently fully autonomous, it requires relatively little input from the user, though expert level knowledge is still required to interpret analysis and provide accurate seed mechanisms and training data. Future work towards the rapid development of such models should focus on the provision of accurate reaction rates and thermochemistry for relevant functional groups and their

inclusion into relevant AMG databases. This would not only improve predictive accuracy but should also minimize the need for mechanism post-processing.

ACKNOWLEDGMENTS

The authors would like to thank all members of the Sus-LABB (Sustainable Low Cost Lignocellulosic Advanced Biofuel Blends) group at the University of Leeds and Trinity College Dublin for their input and support during this project. Appreciation is also extended to the RMG group for providing easy, open access to the RMG software, documentation, and additional resources. This research was funded by the EPSRC grant EP/T033088/1.

DATA AVAILABILITY STATEMENT

No new experimental data is produced in this study and all the materials required to replicate the computational work are provided in supplementary materials or the literature.

ORCID

Christian A. Michelbach  <https://orcid.org/0000-0001-6448-2705>

REFERENCES

1. Masson-Delmotte V, Zhai P, Pörtner HO, et al. Global Warming of 1.5°C: IPCC Special Report on Impacts of Global Warming of 1.5°C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty, IPCC; 2018.
2. Shukla PR, Skea J, Slade R, et al. others, Climate change 2022: Mitigation of climate change, Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Vol. 10; 2022. 9781009157926.
3. IEA. 2022, World Energy Outlook 2022, Paris; 2022.
4. Hanssen SV, Daioglou V, Steinmann ZJN, Doelman JC, Van Vuuren DP, Huijbregts MAJ. The climate change mitigation potential of bioenergy with carbon capture and storage. *Nat Clim Chang*. 2020;10:1023-1029.
5. Leitner W, Klankermayer J, Pischinger S, Pitsch H, Kohse-Höinghaus K. Advanced biofuels and beyond: chemistry solutions for propulsion and production. *Angew Chem Int Ed*. 2017;56:5412-5452.
6. Field JL, Richard TL, Smithwick EAH, et al. Robust paths to net greenhouse gas mitigation and negative emissions via advanced biofuels. *Proc Natl Acad Sci*. 2020;117:21968-21977.
7. Lange J-P, van der Heide E, van Buijtenen J, Price R. Furfural—a promising platform for lignocellulosic biofuels. *ChemSusChem*. 2012;5:150-166.
8. Howard MS, Issayev G, Naser N, Sarathy SM, Farooq A, Dooley S. Ethanol gasoline, a lignocellulosic advanced biofuel. *Sustainable Energy Fuels*. 2019;3:409-421.
9. Michelbach CA, Tomlin AS. Influence of iso-butanol blending with a reference gasoline and its surrogate on spark-ignition engine performance. *Energy Fuels*. 2021;35:19665-19688.
10. Tran L-S, Herbinet O, Li Y, et al. Low-temperature gas-phase oxidation of diethyl ether: fuel reactivity and fuel-specific products. *Proc Combust Inst*. 2019;37:511-519.
11. Wagon SW, Thion S, Nilsson EJK, et al. Experimental and modeling studies of a biofuel surrogate compound: laminar burning velocities and jet-stirred reactor measurements of anisole. *Combust Flame*. 2018;189:325-336.
12. Cai L, Uygun Y, Togbé C, et al. An experimental and modeling study of n-octanol combustion. *Proc Combust Inst*. 2015;35:419-427.
13. Howard MS, Issayev G, Naser N, Sarathy SM, Farooq A, Dooley S. Correction: ethanol gasoline, a lignocellulosic advanced biofuel. *Sustainable Energy Fuels*. 2022;6:3080-3083.
14. Van de Vijver R, Vandewiele NM, Bhoorasingh PL, et al. Automatic mechanism and kinetic model generation for gas- and solution-phase processes: a perspective on best practices, recent advances, and future challenges. *Int J Chem Kinet*. 2015;47:199-231.
15. Venkatasubramanian V. The promise of artificial intelligence in chemical engineering: is it here, finally? *AIChE J*. 2019;65:466-478.
16. Dobbelaere MR, Plehiers PP, Van de Vijver R, Stevens CV, Van Geem KM. Machine learning in chemical engineering: strengths, weaknesses, opportunities, and threats. *Engineering*. 2021;7:1201-1211.
17. Lee JH, Shin J, Realf MJ. Machine learning: overview of the recent progresses and implications for the process systems engineering field. *Comput Chem Eng*. 2018;114:111-121.
18. Broadbelt LJ, Pfaendtner J. Lexicography of kinetic modeling of complex reaction networks. *AIChE J*. 2005;51:2112-2121.
19. Liu M, Grinberg Dana A, Johnson M. Reaction mechanism generator v3.0: advances in automatic mechanism generation. *J Chem Inf Model*. 2021;61:2686-2696.
20. Thion S, Togbé C, Serinyel Z, Dayma G, Dagaut P. A chemical kinetic study of the oxidation of dibutyl-ether in a jet-stirred reactor. *Combust Flame*. 2017;185:4-15.
21. Galano A, Alvarez-Idaboy JR, Bravo-Pérez G, Ruiz-Santoyo ME. Gas phase reactions of C1–C4 alcohols with the OH radical: a quantum mechanical approach. *Phys Chem Chem Phys*. 2002;4:4648-4662.
22. Rotavera B, Taatjes CA. Influence of functional groups on low-temperature combustion chemistry of biofuels. *Prog Energy Combust Sci*. 2021;86:100925.
23. Bugler J, Marks B, Mathieu O, et al. An ignition delay time and chemical kinetic modeling study of the pentane isomers. *Combust Flame*. 2016;163:138-156.
24. Zhou C-W, Simmie JM, Curran HJ. An ab initio /Rice-Ramsperger-Kassel-Marcus study of the hydrogen-abstraction reactions of methyl ethers, H 3 COCH 3–x (CH 3) x, x = 0–2, by ·OH; mechanism and kinetics. *Phys Chem Chem Phys*. 2010;12:7221-7233.
25. Ranzi E, Dente M, Goldaniga A, Bozzano G, Faravelli T. Lumping procedures in detailed kinetic modeling of gasification, pyrolysis, partial oxidation and combustion of hydrocarbon mixtures. *Prog Energy Combust Sci*. 2001;27:99-139.
26. Broadbelt LJ, Stark SM, Klein MT. Termination of computer-generated reaction mechanisms: species rank-based convergence criterion. *Ind Eng Chem Res*. 1995;34:2566-2573.
27. De Witt MJ, Dooling DJ, Broadbelt LJ. Computer generation of reaction mechanisms using quantitative rate information: application to long-chain hydrocarbon pyrolysis. *Ind Eng Chem Res*. 2000;39:2228-2237.

28. Vandewiele NM, Van Geem KM, Reyniers M-F, Marin GB. Genesys: kinetic model construction using chemo-informatics. *Chem Eng J*. 2012;207–208:526–538.
29. Gao CW, Allen JW, Green WH, West RH. Reaction Mechanism Generator: automatic construction of chemical kinetic mechanisms. *Comput Phys Commun*. 2016;203:212–225.
30. Benson SW, Buss JH. Additivity rules for the estimation of molecular properties. thermodynamic properties. *J Chem Phys*. 1958;29:546–572.
31. Sabbe MK, Saeys M, Reyniers M-F, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *J Phys Chem A*. 2005;109:7466–7480.
32. Vandeputte AG, Sabbe MK, Reyniers M-F, Marin GB. Modeling the gas-phase thermochemistry of organosulfur compounds. *Chem – Eur J*. 2011;17:7656–7673.
33. Sabbe MK, De Vleschouwer F, Reyniers M-F, Waroquier M, Marin GB. First principles based group additive values for the gas phase standard entropy and heat capacity of hydrocarbons and hydrocarbon radicals. *J Phys Chem A*. 2008;112:12235–12251.
34. Paraskevas PD, Sabbe MK, Reyniers M-F, Papayannakos N, Marin GB. Group additive values for the gas-phase standard enthalpy of formation, entropy and heat capacity of oxygenates. *Chem – Eur J*. 2013;19:16431–16452.
35. Wang T, Yalamanchi KK, Bai X, et al. Computational thermochemistry of oxygenated polycyclic aromatic hydrocarbons and relevant radicals. *Combust Flame*. 2023;247:112484.
36. Lay TH, Bozzelli JW, Dean AM, Ritter ER. Hydrogen atom bond increments for calculation of thermodynamic properties of hydrocarbon radical species. *J Phys Chem*. 1995;99:14514–14527.
37. vom Lehn F, Cai L, Pitsch H. Impact of thermochemistry on optimized kinetic model predictions: auto-ignition of diethyl ether. *Combust Flame*. 2019;210:454–466.
38. vom Lehn F, Cai L, Pitsch H. Investigating the impacts of thermochemical group additivity values on kinetic model predictions through sensitivity and uncertainty analyses. *Combust Flame*. 2020;213:394–408.
39. Joback KG. *A unified approach to physical property estimation using multivariate statistical techniques*, Thesis. Massachusetts Institute of Technology; 1984.
40. Joback KG, Reid RC. Estimation of pure-component properties from group-contributions. *Chem Eng Commun*. 1987;57:233–243.
41. Tee LS, Gotoh S, Stewart WE. Molecular parameters for normal fluids. Lennard-Jones 12–6 potential. *Ind Eng Chem Fund*. 1966;5:356–363.
42. Ratkiewicz A, Truong TN. Automated mechanism generation: from symbolic calculation to complex chemistry. *Int J Quantum Chem*. 2006;106:244–255.
43. Green WH. In: Marin GB, ed. *Predictive Kinetics: A New Approach for the 21st Century*. *Advances in Chemical Engineering*. Academic Press; 2007:1–313.
44. Blurock E, Battin-Leclerc F, Faravelli T, Green WH. Automatic generation of detailed mechanisms. In: Battin-Leclerc F, Simmie JM, Blurock E, eds. *Cleaner Combustion: Developing Detailed Chemical Kinetic Models*. Springer; 2013:59–92.
45. Battin-Leclerc F, Glaude PA, Warth V, Fournet R, Scacchi G, Côme GM. Computer tools for modelling the chemical phenomena related to combustion. *Chem Eng Sci*. 2000;55:2883–2893.
46. Miyoshi A, KUCRS—Detailed kinetic mechanism generator for versatile fuel components and mixtures, Proceedings of the 8th International Conference on Modeling and Diagnostics for Advanced Engine Systems, COMODIA 2012. 2012:116–121.
47. De Bruycker R, Herbinet O, Carstensen H-H, Battin-Leclerc F, Van Geem KM. Understanding the reactivity of unsaturated alcohols: experimental and kinetic modeling study of the pyrolysis and oxidation of 3-methyl-2-butenol and 3-methyl-3-butenol. *Combust Flame*. 2016;171:237–251.
48. Hakka MH, Bennadji H, Biet J, et al. Oxidation of methyl and ethyl butanoates. *Int J Chem Kinet*. 2010;42:226–252.
49. Sakai Y, Herzler J, Werler M, Schulz C, Fikri M. A quantum chemical and kinetics modeling study on the autoignition mechanism of diethyl ether. *Proc Combust Inst*. 2017;36:195–202.
50. Curran HJ, Gaffuri P, Pitz WJ, Westbrook CK. A comprehensive modeling study of iso-octane oxidation. *Combust Flame*. 2002;129:253–280.
51. Zhou C-W, Simmie JM, Curran HJ. Rate constants for hydrogen-abstraction by O'H from n-butanol. *Combust Flame*. 2011;158:726–731.
52. Zyada A, Samimi-Abianeh O. Ethanol kinetic model development and validation at wide ranges of mixture temperatures, pressures, and equivalence ratios. *Energy Fuels*. 2019;33:7791–7804.
53. Geem KMV, Reyniers M-F, Marin GB, Song J, Green WH, Matheu DM. Automatic reaction network generation using RMG for steam cracking of n-hexane. *AIChE J*. 2006;52:718–730.
54. Harper MR, Van Geem KM, Pyl SP, Marin GB, Green WH. Comprehensive reaction mechanism for n-butanol pyrolysis and combustion. *Combust Flame*. 2011;158:16–41.
55. Harper MR, Van Geem KM, Pyl SP, Merchant SS, Marin GB, Green WH. Erratum to “Comprehensive reaction mechanism for n-butanol pyrolysis and combustion” [*Combust. Flame* 158 (2011) 16–41]. *Combust Flame*. 2011;158:2075.
56. Allen JW, Scheer AM, Gao CW, et al. A coordinated investigation of the combustion chemistry of diisopropyl ketone, a prototype for biofuels produced by endophytic fungi. *Combust Flame*. 2014;161:711–724.
57. Green WH, Allen JW, Buesser BA, et al. RMG-Java—Reaction Mechanism Generator v3.3; 2011. <http://rmg.sourceforge.net/>
58. Dong X, Pio G, Arafin F, et al. Butyl acetate pyrolysis and combustion chemistry: mechanism generation and shock tube experiments. *J Phys Chem A*. 2023;127:3231–3245.
59. Bailey B, Eberhardt J, Goguen S, Erwin J, Diethyl Ether (DEE) as a Renewable Diesel Fuel, In: *International Fuels & Lubricants Meeting & Exposition*, SAE International; 1997.
60. de C e Silva DF. DME as Alternative Diesel Fuel: overview. *2006 SAE Brasil Congress and Exhibit*. SAE International; 2006.
61. Hoppe F, Heuser B, Thewes M, et al. Tailor-made fuels for future engine concepts. *Int J Engine Res*. 2016;17:16–27.
62. Rakopoulos DC, Rakopoulos CD, Giakoumis EG, Dimaratos AM. Characteristics of performance and emissions in high-speed direct injection diesel engine fueled with diethyl ether/diesel fuel blends. *Energy*. 2012;43:214–224.
63. Raspolli Galletti AM, Antonetti C, Fulignati S, Licursi D. Direct alcoholysis of carbohydrate precursors and real cellulosic biomasses to alkyl levulinates: a critical review. *Catalysts*. 2020;10:1221.

64. Zhu W, Chang C, Ma C, Du F. Kinetics of glucose ethanolysis catalyzed by extremely low sulfuric acid in ethanol medium. *Chin J Chem Eng.* 2014;22:238-242.
65. Sivasankaralingam V, Raman V, Mubarak Ali MJ, et al. Experimental and numerical investigation of ethanol/diethyl ether mixtures in a CI engine. In: SAE 2016 International Powertrains, Fuels & Lubricants Meeting, SAE International; 2016.
66. Paul A, Panua RS, Debroy D, Bose PK. Effect of diethyl ether and ethanol on performance, combustion, and emission of single-cylinder compression ignition engine. *Int J Ambient Energy.* 2017;38:2-13.
67. Issayev G, Mani Sarathy S, Farooq A. Autoignition of diethyl ether and a diethyl ether/ethanol blend. *Fuel.* 2020;279:118553.
68. Qi DH, Chen H, Geng LM, Bian YZ. Effect of diethyl ether and ethanol additives on the combustion and emission characteristics of biodiesel-diesel blended fuel engine. *Renewable Energy.* 2011;36:1252-1258.
69. Rakopoulos DC, Rakopoulos CD, Giakoumis EG, Dimaratos AM. Studying combustion and cyclic irregularity of diethyl ether as supplement fuel in diesel engine. *Fuel.* 2013;109:325-335.
70. Uygun Y. Ignition studies of undiluted diethyl ether in a high-pressure shock tube. *Combust Flame.* 2018;194:396-409.
71. Werler M, Cancino LR, Schiessl R, Maas U, Schulz C, Fikri M. Ignition delay times of diethyl ether measured in a high-pressure shock tube and a rapid compression machine. *Proc Combust Inst.* 2015;35:259-266.
72. Yasunaga K, Gillespie F, Simmie JM, et al. A multiple shock tube and chemical kinetic modeling study of diethyl ether pyrolysis and oxidation. *J Phys Chem A.* 2010;114:9098-9109.
73. Zhang Z, Hu E, Peng C, Huang Z. Experimental and kinetic study on ignition delay times of diethyl ether. *SAE Int J Fuels Lubr.* 2015;8:111-118.
74. Belhadj N, Benoit R, Lailliau M, Glasziou V, Dagaut P. Oxidation of diethyl ether: extensive characterization of products formed at low temperature using high resolution mass spectrometry. *Combust Flame.* 2021;228:340-350.
75. Liu X, Zhang Q, Ito S, Wada Y. Oxidation characteristics and products of five ethers at low temperature. *Fuel.* 2016;165:513-525.
76. Sela P, Sakai Y, Choi HS, et al. High-temperature unimolecular decomposition of diethyl ether: shock-tube and theory studies. *J Phys Chem A.* 2019;123:6813-6827.
77. Serinyel Z, Lailliau M, Thion S, Dayma G, Dagaut P. An experimental chemical kinetic study of the oxidation of diethyl ether in a jet-stirred reactor and comprehensive modeling. *Combust Flame.* 2018;193:453-462.
78. Vin N, Herbinet O, Battin-Leclerc F. Diethyl ether pyrolysis study in a jet-stirred reactor. *J Anal Appl Pyrolysis.* 2016;121:173-176.
79. Di Y, Huang Z, Zhang N, Zheng B, Wu X, Zhang Z. Measurement of laminar burning velocities and Markstein lengths for diethyl ether-air mixtures at different initial pressure and temperature. *Energy Fuels.* 2009;23:2490-2497.
80. Gillespie F, Metcalfe WK, Dirrenberger P, et al. Measurements of flat-flame velocities of diethyl ether in air. *Energy.* 2012;43:140-145.
81. Hashimoto J, Tanoue K, Taide N, Nouno Y. Extinction limits and flame structures of 1-butanol and diethyl ether non-premixed flames. *Proc Combust Inst.* 2015;35:973-980.
82. Tran L-S, Pieper J, Carstensen H-H, et al. Experimental and kinetic modeling study of diethyl ether flames. *Proc Combust Inst.* 2017;36:1165-1173.
83. Zhang N, Di Y, Huang Z, Zhang Z. Flame instability analysis of diethyl ether-air premixed mixtures at elevated pressures. *Chin Sci Bull.* 2010;55:314-320.
84. Zhang N, Di Y, Huang Z, Zheng B, Zhang Z. Experimental study on combustion characteristics of N₂-diluted diethyl ether-air mixtures. *Energy Fuels.* 2009;23:5798-5805.
85. Sakai Y, Ando H, Chakravarty HK, Pitsch H, Fernandes RX. A computational study on the kinetics of unimolecular reactions of ethoxyethylperoxy radicals employing CTST and VTST. *Proc Combust Inst.* 2015;35:161-169.
86. Danilack AD, Klippenstein SJ, Georgievskii Y, Goldsmith CF. Low-temperature oxidation of diethyl ether: reactions of hot radicals across coupled potential energy surfaces. *Proc Combust Inst.* 2021;38:671-679.
87. Adler TB, Werner H-J, Manby FR. Local explicitly correlated second-order perturbation theory for the accurate treatment of large molecules. *J Chem Phys.* 2009;130:054106.
88. Adler TB, Knizia G, Werner H-J. A simple and efficient CCSD(T)-F12 approximation. *J Chem Phys.* 2007;127:221106.
89. Georgievskii Y, Miller JA, Burke MP, Klippenstein SJ. Reformulation and solution of the master equation for multiple-well chemical reactions. *J Phys Chem A.* 2013;117:12146-12154.
90. Danilack AD, Mulvihill CR, Klippenstein SJ, Goldsmith CF. Diastereomers and low-temperature oxidation. *J Phys Chem A.* 2021;125:8064-8073.
91. Susnow RG, Dean AM, Green WH, Peczak P, Broadbelt LJ. Rate-based construction of kinetic models for complex systems. *J Phys Chem A.* 1997;101:3731-3740.
92. Johnson MS, Dong X, Grinberg Dana A, et al. RMG database for chemical property prediction. *J Chem Inf Model.* 2022;62:4906-4915.
93. Duan Y, Monge-Palacios M, Grajales-Gonzalez E, Han D, Sarathy SM. Diethyl ether oxidation: revisiting the kinetics of the intramolecular hydrogen abstraction reactions of its primary and secondary peroxy radicals. *Fuel.* 2022;326:125046.
94. Luo Y-R. *Handbook of Bond Dissociation Energies in Organic Compounds.* CRC Press; 2002.
95. Curran HJ. Developing detailed chemical kinetic mechanisms for fuel combustion. *Proc Combust Inst.* 2019;37:57-81.
96. Lai L, Khanniche S, Green WH. Thermochemistry and group additivity values for fused two-ring species and radicals. *J Phys Chem A.* 2019;123:3418-3428.
97. Han K, Jamal A, Grambow CA, Buras ZJ, Green WH. An extended group additivity method for polycyclic thermochemistry estimation. *Int J Chem Kinet.* 2018;50:294-303.
98. Farina DS Jr, Sirumalla SK, Mazeau EJ, West RH. Extensive high-accuracy thermochemistry and group additivity values for halocarbon combustion modeling. *Ind Eng Chem Res.* 2021;60:15492-15501.
99. Ince A, Carstensen H-H, Sabbe M, Reyniers M-F, Marin GB. Group additive modeling of substituent effects in monocyclic aromatic hydrocarbon radicals. *AIChE J.* 2017;63:2089-2106.
100. Ince A, Carstensen H-H, Reyniers M-F, Marin GB. First-principles based group additivity values for thermochemical properties of substituted aromatic compounds. *AIChE J.* 2015;61:3858-3870.

101. Sumathi R, Green WH. Thermodynamic properties of ketenes: group additivity values from quantum chemical calculations. *J Phys Chem A*. 2002;106:7937-7949.
102. Wong BM, Matheu DM, Green WH. Temperature and molecular size dependence of the high-pressure limit. *J Phys Chem A*. 2003;107:6206-6211.
103. Allen JW, Goldsmith CF, Green WH. Automatic estimation of pressure-dependent rate coefficients. *Phys Chem Chem Phys*. 2011;14:1131-1155.
104. Chang AY, Bozzelli JW, Dean AM. Kinetic analysis of complex chemical activation and unimolecular dissociation reactions using QRRK theory and the modified strong collision approximation. *Zeitschrift Für Physikalische Chemie*. 2000;214:1533.
105. Khanshan FS, West RH. Using Reaction Mechanism Generator (RMG) to build detailed kinetic model of biofuels. In: 2013 AIChE Annual Meeting, San Francisco, California, 2013.
106. Burke MP, Chaos M, Ju Y, Dryer FL, Klippenstein SJ. Comprehensive H₂/O₂ kinetic model for high-pressure combustion. *Int J Chem Kinet*. 2012;44:444-474.
107. Li X, Jasper AW, Zádor J, Miller JA, Klippenstein SJ. Theoretical kinetics of O + C₂H₄. *Proc Combust Inst*. 2017;36:219-227.
108. Li Y, Zhou C-W, Somers KP, Zhang K, Curran HJ. The oxidation of 2-butene: a high pressure ignition delay, kinetic modeling study and reactivity comparison with isobutene and 1-butene. *Proc Combust Inst*. 2017;36:403-411.
109. Glarborg P, Miller JA, Ruscic B, Klippenstein SJ. Modeling nitrogen chemistry in combustion. *Prog Energy Combust Sci*. 2018;67:31-68.
110. Wang Z, Yan C, Mei B, Lin Y, Ju Y. Study of low- and intermediate-temperature oxidation kinetics of diethyl ether in a supercritical pressure jet-stirred reactor. *J Phys Chem A*. 2023;127:506-516.
111. Michelbach C, Tomlin A. An experimental and kinetic modeling study of the ignition delay and heat release characteristics of a five component gasoline surrogate and its blends with isobutanol within a rapid compression machine. *Int J Chem Kinet*. 2021;53:787-808.
112. Van de Vijver R, Zádor J. KinBot: automated stationary point search on potential energy surfaces. *Comput Phys Commun*. 2020;248:106947.
113. Goldsmith CF, Magoon GR, Green WH. Database of small molecule thermochemistry for combustion. *J Phys Chem A*. 2012;116:9033-9057.
114. Hughes KJ, Griffiths JF, Fairweather M, Tomlin AS. Evaluation of models for the low temperature combustion of alkanes through interpretation of pressure-temperature ignition diagrams. *Phys Chem Chem Phys*. 2006;8:3197-3210.
115. Johnson MS, Nimlos MR, Ninnemann E, et al. Oxidation and pyrolysis of methyl propyl ether. *Int J Chem Kinet*. 2021;53:915-938.
116. Ye L, Zhang L, Qi F. Ab initio kinetics on low temperature oxidation of iso-pentane: the first oxygen addition. *Combust Flame*. 2018;190:119-132.
117. Mendes J, Zhou C-W, Curran HJ. Theoretical and kinetic study of the reactions of ketones with HO₂ radicals. Part I: abstraction reaction channels. *J Phys Chem A*. 2013;117:4515-4525.
118. Zhou C-W, Simmie JM, Curran HJ. Ab initio and kinetic study of the reaction of ketones with OH for T = 500–2000 K. Part I: hydrogen-abstraction from H₃CC(O)CH₃-x(CH₃)_x, x = 0 → 2. *Phys Chem Chem Phys*. 2011;13:11175-11192.
119. Mendes J, Zhou C-W, Curran HJ. Theoretical and kinetic study of the hydrogen atom abstraction reactions of esters with HO₂ radicals. *J Phys Chem A*. 2013;117:14006-14018.
120. Mendes J, Zhou C-W, Curran HJ. Theoretical study of the rate constants for the hydrogen atom abstraction reactions of esters with •OH radicals. *J Phys Chem A*. 2014;118:4889-4899.
121. Tan T, Yang X, Ju Y, Carter EA. Ab initio kinetics studies of hydrogen atom abstraction from methyl propanoate. *Phys Chem Chem Phys*. 2016;18:4594-4607.
122. Tan T, Yang X, Krauter CM, Ju Y, Carter EA. Ab initio kinetics of hydrogen abstraction from methyl acetate by hydrogen, methyl, oxygen, hydroxyl, and hydroperoxy radicals. *J Phys Chem A*. 2015;119:6377-6390.
123. Thion S, Zaras AM, Szóri M, Dagaut P. Theoretical kinetic study for methyl levulinate: oxidation by OH and CH₃ radicals and further unimolecular decomposition pathways. *Phys Chem Chem Phys*. 2015;17:23384-23391.
124. Wang Q-D, Wang X-J, Liu Z-W, Kang G-J. Theoretical and kinetic study of the hydrogen atom abstraction reactions of ethyl esters with hydrogen radicals. *Chem Phys Lett*. 2014;616-617:109-114.
125. Thion S, Diévert P, Van Cauwenberghe P, Dayma G, Serinyel Z, Dagaut P. An experimental study in a jet-stirred reactor and a comprehensive kinetic mechanism for the oxidation of methyl ethyl ketone. *Proc Combust Inst*. 2017;36:459-467.
126. Hu E, Chen Y, Zhang Z, Chen J-Y, Huang Z. Ab initio calculation and kinetic modeling study of diethyl ether ignition with application toward a skeletal mechanism for CI engine modeling. *Fuel*. 2017;209:509-520.
127. Mendes J, Zhou C-W, Curran HJ. Rate constant calculations of H-atom abstraction reactions from ethers by HO₂ radicals. *J Phys Chem A*. 2014;118:1300-1308.
128. Cavallotti C, Pelucchi M, Frassoldati A. Analysis of acetic acid gas phase reactivity: rate constant estimation and kinetic simulations. *Proc Combust Inst*. 2019;37:539-546.
129. Zhou C-W, Simmie JM, Curran HJ. Rate constants for hydrogen abstraction by HO₂ from n-butanol. *Int J Chem Kinet*. 2012;44:155-164.
130. Power J, Somers KP, Nagaraja SS, Curran HJ. Hierarchical study of the reactions of hydrogen atoms with alkenes: a theoretical study of the reactions of hydrogen atoms with C₂–C₄ alkenes. *J Phys Chem A*. 2021;125:5124-5145.
131. Zhou C-W, Simmie JM, Somers KP, Goldsmith CF, Curran HJ. Chemical kinetics of hydrogen atom abstraction from allylic sites by 3O₂; implications for combustion modeling and simulation. *J Phys Chem A*. 2017;121:1890-1899.
132. Yang F, Deng F, Pan Y, Zhang Y, Tang C, Huang Z. Kinetics of hydrogen abstraction and addition reactions of 3-hexene by OH radicals. *J Phys Chem A*. 2017;121:1877-1889.
133. Le XT, Mai TVT, Ratkiewicz A, Huynh LK. Mechanism and kinetics of low-temperature oxidation of a biodiesel surrogate: methyl propanoate radicals with oxygen molecule. *J Phys Chem A*. 2015;119:3689-3703.
134. Villano SM, Huynh LK, Carstensen H-H, Dean AM. High-pressure rate rules for alkyl + O₂ reactions. 1. The dissociation, concerted elimination, and isomerization channels of the alkyl peroxy radical. *J Phys Chem A*. 2011;115:13425-13442.
135. Ning H, Gong C, Li Z, Li X. Pressure-dependent kinetics of initial reactions in iso-octane pyrolysis. *J Phys Chem A*. 2015;119:4093-4107.

136. Zhao L, Yang T, Kaiser RI, et al. Combined experimental and computational study on the unimolecular decomposition of JP-8 jet fuel surrogates. I. n-Decane (n-C₁₀H₂₂). *J Phys Chem A*. 2017;121:1261-1280.
137. Tranter RS, Klippenstein SJ, Harding LB, Giri BR, Yang X, Kiefer JH. Experimental and theoretical investigation of the self-reaction of phenyl radicals. *J Phys Chem A*. 2010;114:8240-8261.
138. Zhang Y, El-Merhubi H, Lefort B, Le Moyne L, Curran HJ, Kéromnès A. Probing the low-temperature chemistry of ethanol via the addition of dimethyl ether. *Combust Flame*. 2018;190:74-86.
139. Chen D, Wang K, Wang H. Violation of collision limit in recently published reaction models. *Combust Flame*. 2017;186:208-210.
140. Da Silva G, Bozzelli JW, Liang L, Farrell JT. Ethanol oxidation: kinetics of the α -hydroxyethyl radical + O₂ reaction. *J Phys Chem A*. 2009;113:8923-8933.
141. Montgomery JA Jr, Frisch MJ, Ochterski JW, Petersson GA. A complete basis set model chemistry. VII. Use of the minimum population localization method. *J Chem Phys*. 2000;112:6532-6542.
142. Pepiot-Desjardins P, Pitsch H. An efficient error-propagation-based reduction method for large chemical kinetic mechanisms. *Combust Flame*. 2008;154:67-81.
143. Sakai Y, Nakamura H, Sugita T, Tezuka T, Uygun Y. Chemical interpretation on the multi-stage oxidation of diethyl ether. *J Therm Sci*. 2023;32:513-520.
144. Mulvihill CR, Danilack AD, Goldsmith CF, et al. Non-Boltzmann effects in chain branching and pathway branching for diethyl ether oxidation. *Energy Fuels*. 2021;35:17890-17908.
145. Miron L, Chiriac R, Brabec M, Bădescu V. Ignition delay and its influence on the performance of a Diesel engine operating with different Diesel-biodiesel fuels. *Energy Reports*. 2021;7:5483-5494.
146. Krisman A, Mounaïm-Rousselle C, Sivaramakrishnan R, Miller JA, Chen JH. Reference natural gas flames at nominally autoignitive engine-relevant conditions. *Proc Combust Inst*. 2019;37:1631-1638.
147. Griffiths JF, Hughes KJ, Porter R. The role and rate of hydrogen peroxide decomposition during hydrocarbon two-stage autoignition. *Proc Combust Inst*. 2005;30:1083-1091.
148. Vuilleumier D, Selim H, Dibble R, Sarathy M. Exploration of heat release in a homogeneous charge compression ignition engine with primary reference fuels; 2013: 2013-01-2622.
149. Dewey NS, Rotavera B. Reaction mechanisms of alkyloxiranes for combustion modeling. *Combust Flame*. 2023;252:112753.
150. Hartness SW, Rotavera B. Dependence of biofuel ignition chemistry on OH-initiated branching fractions. *Front Mech Eng*. 2021;7:718598.
151. Jena A, Sonawane U, Agarwal AK. Partially premixed combustion of diesel-di-ethyl ether blends in light-duty commercial engine. *Fuel*. 2023;345:128197.
152. Singh AP, Sonawane U, Agarwal AK. Methanol/ethanol/butanol-gasoline blends use in transportation engine—part 1: combustion, emissions, and performance study. *J Energy Res Technol*. 2022;144:102304.
153. Mack JH, Flowers DL, Buchholz BA, Dibble RW. Investigation of HCCI combustion of diethyl ether and ethanol mixtures using carbon 14 tracing and numerical simulations. *Proc Combust Inst*. 2005;30:2693-2700.
154. Nagarajan G, Rao AN, Renganarayanan S. Emission and performance characteristics of neat ethanol fuelled DI diesel engine. *Int J Ambient Energy*. 2002;23:149-158.
155. Polat S. An experimental study on combustion, engine performance and exhaust emissions in a HCCI engine fuelled with diethyl ether-ethanol fuel blends. *Fuel Process Technol*. 2016;143:140-150.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Michelbach CA, Tomlin AS. Automatic mechanism generation for the combustion of advanced biofuels: A case study for diethyl ether. *Int J Chem Kinet*. 2023;1-30.
<https://doi.org/10.1002/kin.21705>