Automated Reaction Exploration of Ozonation Processes for Model Olefins in Water

Enric Petrus,[†,‡] Livia A. Hunkeler,[‡] Markus Reiher,[‡] Urs von Gunten,[†,¶] and Thomas B. Hofstetter[*,†,§]

†*Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland*

[1]

‡*Department of Chemistry and Applied Biosciences, ETH Zürich, Vladimir-Prelog-Weg 2, 8093 Zürich, Switzerland*

¶*School of Architecture, Civil and Environmental Engineering (ENAC), École Polythecnique Fédéral de Lausanne (EPFL), 1015 Lausanne, Switzerland*

§*Institute of Biogeochemistry and Pollutant Dynamics (IBP), ETH Zürich, 8092 Zürich, Switzerland*

E-mail: thomas.hofstetter@eawag.ch

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Abstract

The comprehensive evaluation of pollutant abatement during chemical oxidation processes and identification of potentially hazardous transformation products are a fundamental challenge in water and wastewater treatment. Here, we demonstrate how high-throughput computational chemistry enables the elucidation of reaction pathways via automated, quantum-chemistry-based chemical reaction network (CRN) explorations. We evaluated the predictive capabilities of this computational approach using the Software for Chemical Interaction Networks (SCINE) for studying the reactions of ozone with two olefins, ethene and tetramethylethene, in aqueous solution. Following a benchmarking of the quantum chemical methodology for structure optimization and energy calculations, we generated CRNs containing hundreds of compounds and thousands of reactions, identified reaction mechanisms, and evaluated product formation kinetics through microkinetic modeling. These CRN explorations led to the correct reproduction of experimental evidence for mechanisms and products of olefin ozonolysis for reactions of ozone and ethene based solely on defining the reactants and their initial concentrations. The study of reactions of ozone and tetramethylethene also matched experimental data for the main products but revealed consequences of limited exploration depth and shortcomings of the implicit solvation model. We envision that CRN explorations not only offer novel means for predicting pollutant transformation pathways but will also support chemical analysis and the assessment of effects on human and environmental health.

Keywords: water treatment, ozonation, chemical reaction network, quantum chemistry, high-throughput computational chemistry microkinetic modeling, reaction pathway identification

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Introduction

The use of chemical oxidants including ozone, chlorine, and chlorine dioxide in water treatment processes is essential to mitigate waterborne diseases and abate inorganic and organic micropollutants.[1,2] Chemical oxidants have been used successfully for decades for the provision of safe drinking water and, more recently, for wastewater treatment.[2] Their reactions with matrix components and organic pollutants create transformation and disinfection (by)products which can cause adverse effects on human and environmental health.[3–12] Providing efficient disinfection/oxidation while minimizing toxic product formation is particularly challenging for wastewater treatment and potable water reuse purposes where concentrations of dissolved organic matter and organic pollutants are typically much higher.[13–15] Evaluating the efficacy of pollutant abatement during chemical oxidation processes and identifying potentially hazardous transformation products are thus paramount for the implementation of adequate water treatment strategies.[16,17]

To date, there is a large data base on reactions and mechanisms of oxidation reactions, however, comprehensive experimental descriptions of all reaction pathways pertinent to oxidative water treatment would be extremely laborious and experimental evidence remains inherently incomplete.[1] Computational support is therefore required to complement the experimental characterizations of oxidation processes. In fact, chemoinformatic (i.e., data-driven) pathway prediction tools and first-principles quantum chemical calculations have already assisted the experimental elucidation of kinetics and reaction pathways. Previously, empirical reaction rules for reactive functional groups of compounds with unknown reactivity towards oxidants were applied to predict the formation of transformation products.[18,19] Furthermore, quantum chemical descriptors for structure activity relationships were established, allowing access to chemical structures, their Gibbs energies, and barrier heights for the characterization of reaction pathways.[20–25] However, quantum chemical construction of potential energy surfaces (PES) requires significant manual input and guidance, limiting computational analyses to specific reaction coordinates. These circumstances have so far

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

prevented a more expansive, systematic, and possibly predictive evaluation of reactions involving chemical oxidants and organic micropollutants.

These restrictions can now be overcome with high-throughput computational chemistry which offers new avenues to elucidate reaction pathways and mechanisms through chemical reaction network (CRN) explorations.[26–29] Supported by the steady increase of computing power and availability of high performance computing clusters, quantum chemistry based reaction network explorations are developing rapidly and they make extensive use of algorithms for their automated and autonomous execution. Several mechanism-exploration procedures have been implemented in recent years (see compilation in ref [30]). In the Software for Chemical Interaction Networks (SCINE),[26] which we will use in this study, such CRN explorations are carried out in three principle steps. First, the reaction space is explored by searching elementary steps which interconnect the molecular structures of reactants, intermediates, transition states and products. To do so, first-principle heuristics, that is features of the electronic wave function, facilitate the generation of reaction coordinates and thus prediction of reactive sites in molecular structures for uni- and bimolecular reactions. Not only are these algorithms agnostic to the functional groups in the studied molecules, but they also allow for exploring all possible reactive combinations to obtain a comprehensive picture of the reaction network. Second, the minimum and transition-state structures and corresponding energies generated during the initial CRN explorations may require a refinement given that these explorations are typically carried out with fast electronic structure methods such as semipirical ones (e.g., xTB[31,32]) and density functional theory (DFT). These quantum chemical characterizations of the CRN can involve computationally more demanding coupled cluster (CC) methods for electronic energy calculations.[27] Third, the obtained reaction Gibbs energies need to be evaluated modeling experimental conditions in terms of shortest reaction path to determine reaction mechanisms,[33] and microkinetic simulations to predict the concentrations of all the compounds in the CRN.[34] Most recent developments address the need for efficient navigation of CRN explorations and restrict the combinatorial explosion of

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

reaction paths by allowing user-defined interventions (e.g., by expert knowledge, on-the-fly evaluation of concentration fluxes, interactive structure manipulation).[35–37]

The goal of this study was to evaluate how automated CRN explorations can support the elucidation of mechanisms and pathways of reactions of chemical oxidants with organic compounds and provide a conceptual basis for future applications. We focused our study on reactions of ozone with olefins in aqueous solution for purposes of both practical relevance and systematic development of the computational methodology for CRN explorations. Reactions of ozone with ethenes through the 1,3-dipolar cycloaddition (Criegee) mechanism are furthermore prototypical for reactions of electron-rich olefins and aromatic compounds in water treatment and their kinetics and mechanisms are well-documented with experimental data.[20,38,39] Moreover, ozone has biradical and thus multireference character[40] and its computational identification and optimization of molecular structures and predicting the reaction energies may not be accurate with fast DFT methods, as illustrated in studies of the Criegee mechanism.[22] Instead, single-reference coupled cluster methods are recommended,[41,42] though they can be challenging to use given the large number of molecular structures in CRNs. The specific objectives of this study were threefold. (i) We identified the optimal computational methodology for automated CRN explorations with the Chemoton software[30] through systematic evaluation of the Criegee mechanism for ozone and ethene against benchmark studies of Wheeler et al.[41] (ii) We tested the predictive capabilities of automated CRN explorations by investigating two model reactions of ozone with ethene and tetramethylethene, respectively. These processes were selected as benchmarks because they both proceed through the Criegee mechanism but generate different stable products under identical experimental conditions.[38] Here, we performed the CRN evaluation with methods identified in specific objective (i) and evaluated their utility for future automated CRN explorations involving ozone and organic compounds. (iii) Finally, we analysed the outcome for the lowest energy path reaction mechanism between the reactants and the experimentally observed products with Pathfinder[33] as well as through microkinetic modeling.[34]

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Computational Methodology

Quantum Chemical Calculations

Density functional theory and coupled cluster calculations were performed with the ORCA software package version 5.0.3.[43] We also tested three semiempirical methods, namely DFTB3,[44] PM6[45] and GNF2-xTB.[32] We have not considered the application of multireference configuration interaction methods because of the two following limitations. First, molecular orbitals that form the complete active space must be consistent for all the molecular structures along a reaction path, hence new methodologies are currently under development to provide direct orbital selection mapping.[46] Second, multireference methods are notorious for being among the most time-demanding calculations in quantum chemistry.[47] Given that pioneering studies demonstrated that CC succeeded in both optimizing the molecular structures and predicting the reaction energies of the Criegee mechanism for ozone and ethene, we have followed their recommendation.[41,42] DFT structure optimizations were done with Perdew-Becke-Ernzerhof PBE,[48,49] the hybrid PBE0,[50] and the long-range LC-PBE[51] density functionals, with the Ahlrichs def2-TZVP basis set.[52] Grimme's D3 dispersion correction was employed for the PBE and PBE0 calculations.[53,54] Single point calculations were performed with canonical coupled cluster, CCSD(T),[55] and the Domain-based Local Pair Natural Orbital approximation of coupled cluster, DLPNO-CCSD(T)[56] with the correlation consistent basis sets.[57] Because the reaction network contains both open- and closed-shell compounds, we employed the DLPNO approach, instead of the DPNO, to ensure that we can combine their absolute electronic energies when calculating reaction energies. We employed the default protocol in ORCA, where UHF alpha and beta orbitals are transformed to quasi-restricted orbitals, thus removing much of the spin contamination. Solvation in water was introduced using the Conductor-like Polarizable Continuum Model, CPCM, both for structure optimization and single point calculations.[58] Stationary points were characterized with analytic frequency calculations. Gibbs energy corrections were computed using LC-PBE, at 298.15 K and 1

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

atm, and using the ideal gas-rigid rotor-harmonic oscillator (IGRRHO) model. Final Gibbs energies were calculated combining the electronic energy from DLPNO-CCSD(T), and the Gibbs energy corrections from the DFT method. In order to employ molar concentrations in the rate constant, the reference state was modified to 298.15 K and 1.0 M (standard state in solution). Given that the temperature remains constant between both states, this was done by simply increasing the barrier by 7.90 kJ·mol^{-1}.[59]

Reaction Network Exploration

We carried out autonomous reaction network explorations with ethene and ozone as well as tetramethylethene and ozone using the SCINE Chemoton module.[60] An introduction to the terminology, the principles of CRN construction, and the characterization of CRNs is provided in Section S1 of the Supporting Information. Briefly, CRN explorations generate four datasets. (i) A *structure* is a single identity defined by a set of atoms distributed in the three-dimensional space with a determined spin and charge. (ii) A *compound* is a group of *structures* with same number of atoms, charge and spin but different electronic energies. (iii) The *elementary step* represents a rearrangement of chemical bonds and electrons between two *structures* connected via a transition state (TS). Finally, (iv) *reactions* comprise a group of *elementary steps* connecting two *compounds* through a transition state (Figure S1).

To generate transition state guesses for uni- and bimolecular reactions, we employed the algorithm implemented in Chemoton named *Newton Trajectory scan* (NT2), which is inspired by the Artificial Force Induced Reaction (AFIR),[61] and the use of Newtonian forces for computing an effective PES.[62] We restricted the size of the compounds in the two networks as follows. NT2 trials that led to activation barriers greater than 250 kJ·mol^{-1} (according to the method which was used for the exploration) were dismissed, as the subsequent reaction would be unlikely to happen. The activation barrier threshold was set relatively high to account for two phenomena. First, electronic structure methods can overestimate experimental barrier heights. Too low thresholds could lead to reactions being overlooked. Second, the

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

identification of barriers occurs regardless of the direction in which they might be crossed. Endergonic reactions with high barriers can still represent exergonic reactions in the opposite direction. The chemical reaction network of ethene and ozone (CRN-E) contained a maximum of 2 carbon atoms, 4 hydrogen atoms, and 3 oxygen atoms (i.e., the ozonide) to reduce the computational cost of the exploration (see further discussion below and Section S2.2). CRN-T for reactions of tetramethylethene and ozone was made up of a maximum of 6 carbon atoms, 14 hydrogen atoms, and 3 oxygen atoms. This threshold has two hydrogen atoms more than the tetramethylethene-ozonide, in order to allow the search for the relevant transformation product, 2,3-dimethylbutane-2,3-diol. We stopped the explorations once the expected, experimentally confirmed products were found.

All calculations were handled by SCINE Puffin instances[63] and stored and processed in the SCINE Database.[64] These instances interface three additional SCINE packages: ReaDuct,[65] for communicating with ORCA's generation of input and output files, Molassembler,[66] for handling molecular structures, and Utilities[67] which contains a broad variety of functionalities essential for all SCINE modules. We monitored the progress of the reaction exploration with the open-source graphical user interface Heron,[68,69] and the reaction networks can be interactively visualized as standalone HTML files generated with VizChemoton.[70] A detailed description of the parameters employed in the exploration is provided in Section S2.2.

CRN Characterization

CRN-E and CRN-T were characterized with reaction pathway identification and kinetic analyses. An analysis of equilibrium species distributions based on thermodynamic stability was not conducted, as ozonation of olefins is driven by fast kinetics. We employed Pathfinder and its built-in cost function to determine the most favorable reaction mechanisms.[33] For the kinetics analyses we used microkinetic simulations with KiNetX's original Matlab code (see parameters in Section S2.7).[34] Initial concentrations for both the pathway identification

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 ORCID: https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

and kinetic analysis were set to $4.97{\cdot}10^{-5}$ M of ozone, $4.97{\cdot}10^{-3}$ M of the model olefin, and excess of water.

Results and Discussion

Choice of Quantum Chemistry Methodology

Reference Reaction

To determine the adequate quantum chemistry method for the reaction exploration, we evaluated the performance of a broad range of methods, focusing on the first step of the ozonolysis reaction. The 1,3-dipolar cycloaddition of ozone to the double bond of the olefin leads to an ozonide intermediate. Figure 1 shows the mechanism of the ozonolysis, known as the Criegee mechanism,[71] for the two olefins, ethene and tetramethylethene, selected in this study. We used as reference the data reported by Wheeler et al.[41] on the 1,3-dipolar cycloaddition reactions, which consisted of the molecular structures of ozone, ethene, and the TS, as well as the corresponding activation energy (14.2 $kJ{\cdot}mol^{-1}$). Moreover, Figure 1 depicts the expected products for the aqueous ozonolysis of ethene (formaldehyde and α-hydroxymethylhydroperoxide) and tetramethylethene (acetone and hydrogen peroxide). The final products of the ozonolysis and their yields reported in Dowideit and von Sonntag[38] are used in the final section of this study to validate the outcome of the CRN exploration in kinetic simulations.

Transition State Structure Optimization

We tested three semiempirical methods, DFTB3, PM6 and GFN2-xTB, and three versions of the Perdew-Becke-Ernzerhof functional, PBE, PBE0, and LC-PBE, to optimize the transition state structure leading to the ozonide in the first part of the method benchmarking step. We did not employ CC methods to optimize molecular structures because calculations would become too slow. Figure 2 shows the the results of structure and energy benchmarks

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 ORCID: https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

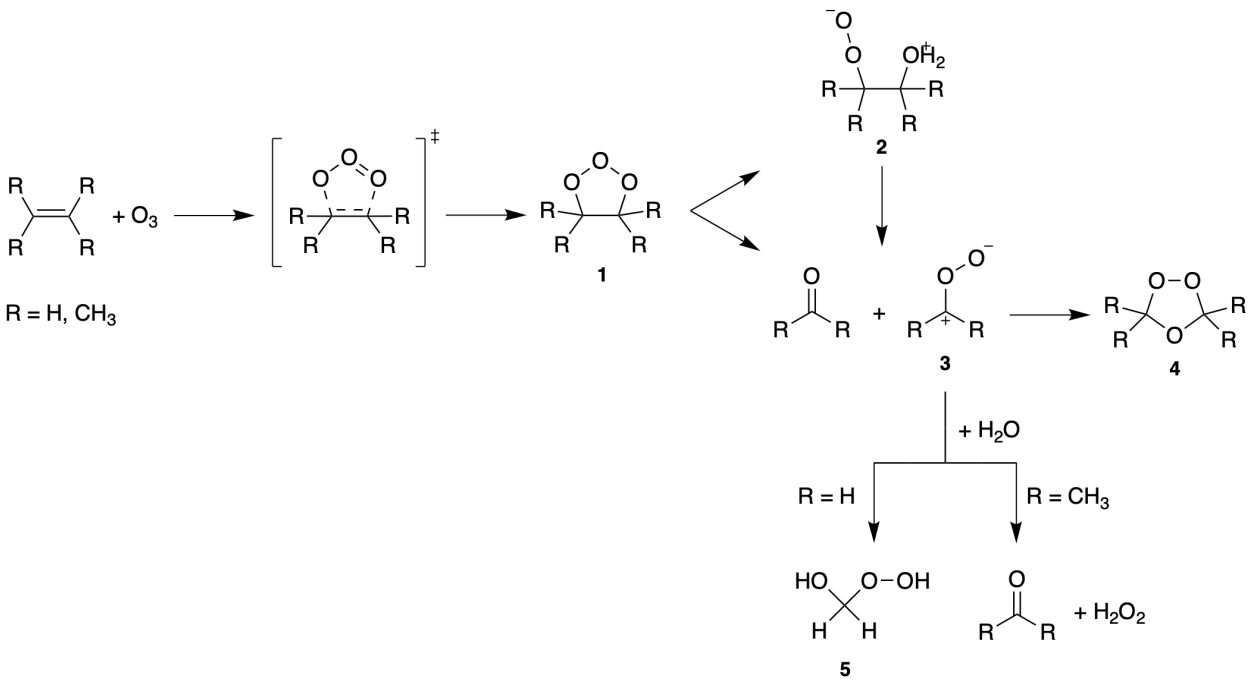


Figure 1: Ozonolysis of ethene and tetramethylethene through the ozonide intermediate (**1**). The Criegee mechanism[71] proceeds either directly with a heterolytic cleavage forming a carbonyl compound and an alkyliumperoxolate (**3**) or via a zwitterionic intermediate (**2**), the latter leading to aldehyde and α-hydroxyl-alkylperoxide (**5**) in presence of water as solvent, and to the Criegee ozonide (**4**) in apolar solvents.[38]

that we conducted to identify the optimal quantum chemistry methodology for ensuing CRN explorations. The data in Figure 2 demonstrate that no single quantum chemistry method simultaneously offers low deviations in molecular structures and electronic energies. Therefore, it is necessary to define a composite approach, where molecular structures are optimized first with one method, and electronic energies are subsequently calculated with a second method.

Figure 2A shows the root mean squared deviation (RMSD) of the TS structures optimized with the aforementioned methods, with respect to the reference TS structure of Wheeler et al.[41] DFTB3 provides a molecular structure with a RMSD above 1.04 Å, followed by PM6 (0.59 Å). GFN2-xTB outperforms the two previous semiempirical methods with a RMSD of 0.20 Å. Neither PBE nor PBE0 succeeded in converging the TS structure, whereas LC-PBE not only locates the TS structure, but also provides the lowest RMSD (0.03 Å). Considering that commonly used RMSD thresholds range from 0.5 to 2.0 Å,[72] only GFN2-xTB and

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

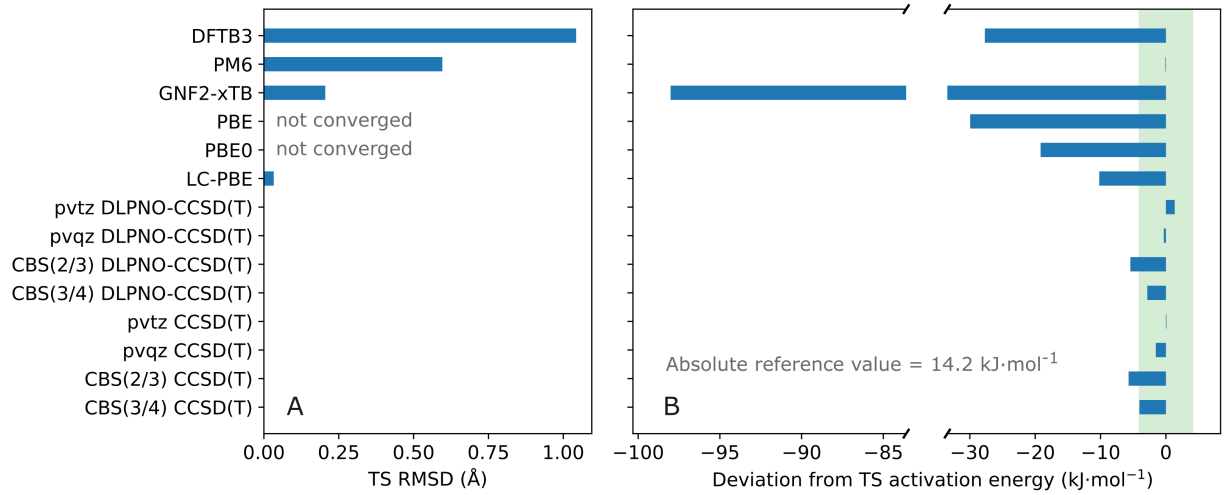


Figure 2: Benchmark of 14 quantum chemistry methods sorted from the fastest (top) to the slowest calculation (bottom). RMSD of the optimized structures are shown with respect to the reference TS structure (panel A) and single point energy deviations relative to the reference activation energy of the ozonide formation (panel B), both provided by of Wheeler et al.[41] (14.2 kJ·mol^{-1}). The green rectangle depicts the range of chemical accuracy of ±4.18 kJ·mol^{-1}.

LC-PBE are meaningful candidates for running the reaction exploration. The former is on average 100 times faster than the latter, however, LC-PBE is approximately 10 times more accurate than GFN2-xTB. The excellent performance of LC-PBE is explained by the accurate electronic structures for zwitterionic compounds predicted with range-separated functionals.[73] This point is particularly relevant since not only does ozone exhibit partly zwitterionic character,[74] but also many products of the reaction network are zwitterions. Considering that standard quantum chemistry methods (PBE and PBE0) failed to locate the TS and that semiempirical methods are persistently inaccurate for complex bond-breaking processes, we employed LC-PBE for finding and optimizing the molecular structures in our CRN explorations.

Accuracy of Transition State Energies

We have evaluated the performance of 14 quantum chemistry methods (3 semiempirical methods, 3 DFT functionals, 4 DLPNO CC, and 4 canonical CC) to predict the activation

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

energy of the ozonide. Figure 2B shows how the electronic energies calculated with the 14 methods, using single point calculations with the reference structures from Wheeler et al. [41], deviate from the reference value of 14.2 kJ·mol^{-1}.[41] Deviations vary substantially for the 3 semiempirical methods from −83.6 kJ·mol^{-1} for GFN2-xTB to 0.08 kJ·mol^{-1} for PM6. We attribute PM6's high accuracy to its parametrization with CC energy data of its non-covalent interactions.[75] The energy deviations of DFT calculations, that is PBE, PBE0 and LC-PBE are −29.9, −19.1, and −10.1 kJ·mol^{-1}, respectively. This energy trend follows the Jacob's ladder principle by which a more precise treatment of the exchange-correlation term enhances the prediction of the electronic energy.[76] However, none of the activation energies calculated with these methods are within the chemical accuracy range of ±4.18 kJ·mol^{-1}.[77] The poor accuracy of TS energy calculations with DFT functionals in this type of reactivity agrees with a previous benchmark study.[22] Only data obtained from DLPNO and canonical CC resulted in reliable TS activation energies, as depicted with a green rectangle in Figure 2B. Despite the fact that DLPNO is an approximation of the canonical CC, the former demonstrates a similar accuracy to the latter, even when pvtz CCSD(T) is approximately three times slower than pvtz DLPNO-CCSD(T) (Figure S4). In fact, the wall clock time of the canonical CC rapidly scales exponentially for molecules with more than 20 atoms, whereas DLPNO CC scales linearly even with molecules up to 1000 atoms.[78] Moreover, complete basis set extrapolations (CBS) shows a higher deviation than the triple and quadruple zeta basis sets. Data in Figure S4 also illustrates that pvqz DLPNO-CCSD(T) is 4.1 times more accurate than pvtz DLPNO-CCSD(T), however, it is also 4.7 times more expensive. Based on this evidence, electronic energies in our CRN explorations were calculated with pvtz DLPNO-CCSD(T).

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Generation of Chemical Reaction Networks for Ozone and Reference Olefins

Exploration Outcomes

The CRN exploration initiated through the reaction of ethene with ozone, with LC-PBE for structures and DLPNO-CCSD(T) for energies, generated 595 compounds and 1350 reactions (Table 1). The serial computing time for the CRN-E exploration, that is the time invested in all the calculations of the CRN exploration on a hypothetical single processor core,[35] was 28500 days (Table 1). Given that computations are run in parallel, the effective total computing time using an average of 250 cores was 114 days.

We assessed the plausibility of structures generated for CRN-E in analogy to the computational benchmarking process with experimental evidence for the detected reaction products.[20,38] Table 1 shows that the exploration for CRN-E found the the key compounds pertinent to the ozonolysis reactions, namely the ozonide, formaldehyde, methyliumperoxolate, α-hydroxymethylhydroperoxide, and hydrogen peroxide. CHEMOTON successfully deduced the intermediates of the 1,3-dipolar cycloaddition mechanism summarized in Figure 1 without any expert intervention. Water, which is essential for the reaction of the alkyliumperoxolate (**3** in Figure 1), was not included at the start of the exploration but generated by CHEMOTON. Lastly, water was found to be involved in 59 reactions, thus demonstrating the consistency of the autonomous exploration procedure.

Representing the complete network of CRN-E in a single yet meaningful Figure is not doable given the large number of reactions and compounds (Table 1). Here, we illustrate the initial steps with the involved compounds, flasks, and transition states (Figures 3 and S7) leading to the ozonide and the reactions of the ozonolysis reaction path (Figures S5 and S6). A complete representation of the CRNs with all compounds and reactions is available in the HTML files generated with the browser-based visualization module VIZCHEMOTON.[70] Note that the identification of a compound *a priori* does not imply relevance. The latter requires

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

the evaluation of the complete reaction network for reaction mechanisms and kinetics as is done below.

Figure 3 depicts the three initial reactions of CRN-E (orange arrows), one association (yellow arrow), five compounds (green circles) and one flask (blue circle). The exploration starts by pushing together ozone and ethene to form flask A1, which was already found by Wheeler et al.[41] as vdW complex. This pre-reactant complex forms three different compounds through three distinct transition states. The three-dimensional structures of the transition states are depicted in the orange boxes of Figure 3. The exploration procedure identified the most critical compound, the ozonide, C3, through reaction R1, in good agreement with experiments[71] and previous theoretical investigations.[23] Figure 3 also depicts a second reaction, R2, which leads to the formation of the known secondary zwitterion

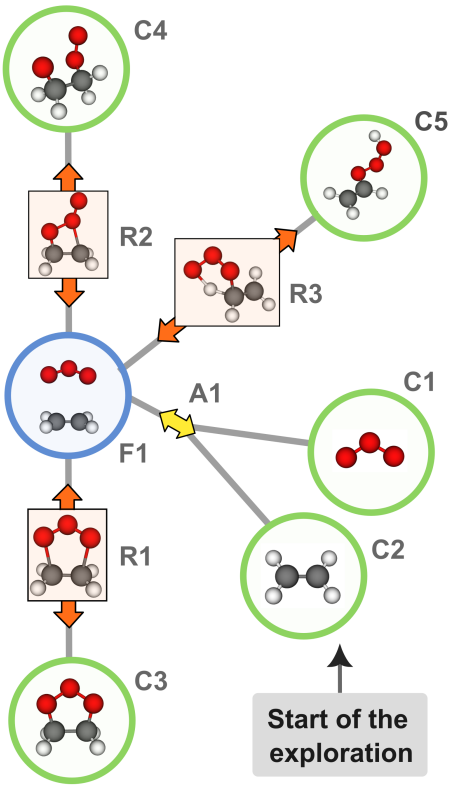


Figure 3: Schematic representation of the initial part of CRN-E constructed by CHEMOTON. Labels of reactions (R) with representations of transition state structures, compounds (C) and flasks (F) named in clockwise order.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 ORCID: https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

Table 1: Features of the three studied chemical reaction network explorations for ozone and ethene (CRN-E) as well as ozone and tetramethylethene (CRN-T and CRN-T-xTB). Total serial computing time, total number of compounds and reactions, number of reactions involving each relevant compound, and classification of reaction types.

	CRN-E	**CRN-T**	**CRN-T-xTB**
Serial computing time (days)	28505	11653	1436
Total number of reactions	1350	441	133426
Total number of compounds	595	388	68721
Compounds	**Number of reactions**		
Ozonide	5	13	31
Formaldehyde	118	5	897
Methyliumperoxolate	29	-	217
α-Hydroxymethylhydroperoxide	6	3	35
Water	59	12	7741
Hydrogen Peroxide	18	9	1872
Acetone	-[a]	12	241
Tetroxane	-[a]	11	-
2-Hydroxy-2-propylhydroperoxide	-[a]	4	12
Reaction classification	**Number of reactions**		
Associations	1710	538	24405
Dissociations	520	158	28028
Bond rearrangements	470	134	81003

[a] not found because of the stoichiometry filter for the CRN-E exploration.

in the Criegee mechanism (**2** in Figure 1).[71] Reaction, R3 shows the formation of the 1-ethenyltrioxidane, C5, which has not been proposed for ozonolysis of ethene.

The exploration for tetramethylethene and ozone, CRN-T, generated 388 compounds and 441 reactions (Table 1). The serial computing time for the CRN-T exploration was 11600 days, and the total computing time, using an average of 250 cores, was reduced to approximately 47 days. We compared some structures generated for CRN-T with the initial steps leading to the tetramethylethene ozonide intermediate in analogy to the CRN-E evaluation. CRN-T contains the key compounds depicted in Figure 1, thus confirming that the product hypothesis is also met for this exploration. The CRN-T exploration also found water as compound. Figure S7 shows an analogous representation of Figure 3 for CRN-T, where the key ozonide compounds and their transition states were successfully identified. However, one experimentally reported transformation product was not found, corresponding

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

to the product of the partial oxidation of tetramethylethene: 2,3-dimethylbutane-2,3-diol.[38] In fact, the reaction exploration did not find any diol compound.

We hypothesize that the CRN-T did not find any diol and, presumably, other intermediates and products because of the inherently vast chemical space. This chemical space would not be accessible within the time frame of our computations with LC-PBE based structure searches. To that end, we ran a separate, reaction exploration starting from tetramethylethene and ozone using GFN2-xTB instead of LC-PBE (abbreviated CRN-T-xTB in Table 1). The computing time running on 250 cores was approximately 6 days (1436 days of serial computing), and during that short period of calculation time, it found 133426 reactions and 68721 compounds. CHEMOTON indeed generated eight different diols (Figure S8), but only after CRN-T-xTB included over 40000 reactions vs. 441 reaction with CRN-T, that is a 100-fold increased depth of the exploration. Further qualitative evidence for the operational limitations of LC-PBE-based CRN explorations was obtained from the consideration of typical intermediates of the aqueous ozone chemistry. CRN-T-xTB found additional important products associated with the general decay of ozone not included in CRN-T such as the hydroxyl radical, the superoxide radical and the ozonide radical (Figure S9). Furthermore, Figure S9 illustrates that singlet and triplet dioxygen, hydrogen peroxide, and water were also present in CRN-T, but involved in a 2- to almost 400-fold lesser frequency than in reactions of CRN-T-xTB. These intermediates did not matter for the reaction of ethene and tetramethylethene with ozone studied here but would do so in ozonation processes with substrates with lower ozone reactivities than olefins. The exploration of the reactions of these ozone decay intermediates, however, was beyond the scope of this study.

Reaction Rate Constants

The DLPNO-CCSD(T) energy calculations for the 1350 reaction in CRN-E enabled us to obtain 2700 forward and backward reaction rates constants for the reactions of 595 compounds in the network. Figure 4A shows the unimodal exponential distribution of reaction

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 ORCID: https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

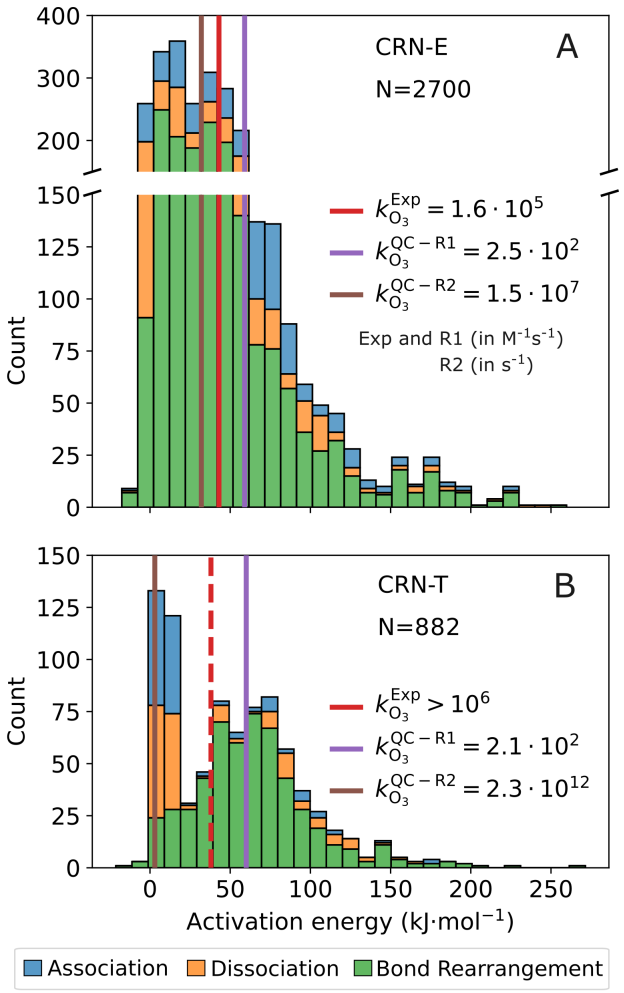


Figure 4: Frequency distribution of activation energies in kJ·mol^{-1} for (A) CRN-E and (B) CRN-T. Bar colors indicate the reaction type, association (blue), dissociation (orange) and bond rearrangement (green). The three color lines depict rate constants in s^{-1} for unimolecular steps and M^{-1}s^{-1} for bimolecular steps. Experimental value in red, $k_{O_3}^{Exp}$, and theoretical values in purple and brown, $k_{O_3}^{QC}$. Reactions QC-R1 and QC-R2 are shown to Figures 3 and S7. Dashed vertical line indicates that only a lower limit of the experimental second-order rate was reported in Dowideit and von Sonntag [38].

rate constants according to their activation energies as confined by the initial boundary conditions of the CRN exploration. The distribution covers a range of approximately 250 kJ·mol^{-1}, which is equivalent to 44 orders of magnitude in reaction rate constants. The peak of the distribution is roughly at 25 kJ·mol^{-1} and the majority of reactions have an energy of <80 kJ·mol^{-1}. Consequently, most reactions have half-lives, $t_{1/2}$, of 12 seconds at room temperature (eq. 1).

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

$$t_{1/2} = \ln(2) \left(\frac{k_B T}{h} \cdot e^{-\Delta G^{\ddagger}/(RT)} \right)^{-1} \quad (1)$$

where k_B, h and R are Boltzmann's, h Planck's, and the gas constant, respectively, $\Delta G^{\ddagger}$ is the activation free energy, and T is the temperature.

We have divided the reaction types in three main groups according to whether the stoichiometry of the reaction (i.e., $\sum \nu_{\text{reactants}} - \sum \nu_{\text{products}}$) is positive (indicative of an association reaction), negative (dissociation) or zero (bond rearrangement). Table 1 shows that CRN-E has 1710 associations, 520 dissociations and 470 bond rearrangements.

A similar ratio between the three reaction types, that is the predominance of association reactions was also found for CRN-T despite the smaller number of compounds and reactions. Figure 4B depicts the bimodal exponential distribution of reaction rate constants according to the activation energies. The two peaks of maximum counts of activation energies in Figure 4B are centered at 8 and 75 kJ·mol^{-1}, respectively. Because the peak at 8 kJ·mol^{-1} is the highest, we deduce that CRN-T contains a relative larger share of fast reactions compared to CRN-E.

While CHEMOTON allows for the autonomous construction of CRNs, the large uncertainty of quantum chemistry methods for predicting absolute values of rate constants still remains.[79] To exemplify issues with absolute rate constants for the CRNs studied here, we have depicted two sets of the experimental and quantum chemical rate constants, $k_{\text{O}_3}^{\text{Exp}}$ and $k_{\text{O}_3}^{\text{QC}}$, for the reaction of ozone with ethene and tetramethylethene, perspectively, in Figure 4. The experimentally determined value of $k_{\text{O}_3}^{\text{Exp}}$ for ethene is shown as a red vertical line, $k_{\text{O}_3}^{\text{Exp}}$ at $1.6 \cdot 10^5$ M^{-1}s^{-1} in Figure 4A.[38] The value of the theoretical $k_{\text{O}_3}^{\text{QC}}$ for reactions R1 and R2 of Figure 3, however, are three and two orders of magnitude different (brown and purple lines in Figure 4A). The mismatch in the absolute value of the rate constants is also observed in the exponential decrease of ozone, as it reacts slower than expected from experimental data (Figure S10). The mismatch between experimental and theoretical data is even more pronounced for tetramethylethene. Because only a lower limit of this second-

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

order rate constant was previously reported,[38] the $k_{O_3}^{Exp}$ is shown as a dashed red vertical line in Figure 4B. $k_{O_3}^{QC}$ corresponds to reactions R1 and R2 of ozone and tetramethylethene (Figure S7) to yield the vdW flask and the ozonide, depicted as purple and a red vertical lines, respectively. Another important consideration is that association reactions suffer from an entropic penalty.[80] On the one hand, the formation of the vdW flask, $k_{O_3}^{QC\text{-}R1}$, is entropically unfavorable, as two reactants, ozone and the olefin, are transformed into one product, the vdW flask. On the other hand, for $k_{O_3}^{QC\text{-}R2}$, that is the reaction of vdW flask to form the ozonide compound (brown vertical line for $k_{O_3}^{R2}$in Figure 4B), the outcome is reversed. These pieces of evidence highlight the fundamental challenge of predicting accurate absolute reaction rate constants with quantum chemistry methods. It thus cannot be the goal of CRN explorations to provide and evaluate individual activation barriers for selected reactions in terms of bimolecular rate constants. Instead, rate constants calculated with the same method (here DLPNO-CCSD(T)), can be used as relative values because they benefit from error cancellation.[59] Relative activation energies are thus consistent and enable one to carry out reaction pathway and mechanisms evaluations as well as microkinetic modeling which we will present in the next section.

Reaction Mechanisms and Kinetics

Reaction Pathway Identification

Figure 5 shows two sets of the 25 most probable reaction mechanisms generated with PATHFINDER for CRN-E and CRN-T in panels A and B, respectively, using initial concentrations from Dowideit and von Sonntag[38] ($4.97 \cdot 10^{-5}$ M of ozone and $4.97 \cdot 10^{-3}$ M of olefin). The most likely mechanism, deduced from a compound cost function that stands for the compound's availability for a reaction,[33] is highlighted in dark blue color, while all other reaction pathways are depicted in grey. The cost has arbitrary units and it is inversely related to the likelihood of the mechanism. In this reaction pathway identification, ozone with ethene and tetramethylethene, respectively, were defined as sources whereas the hy-

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 ORCID: https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

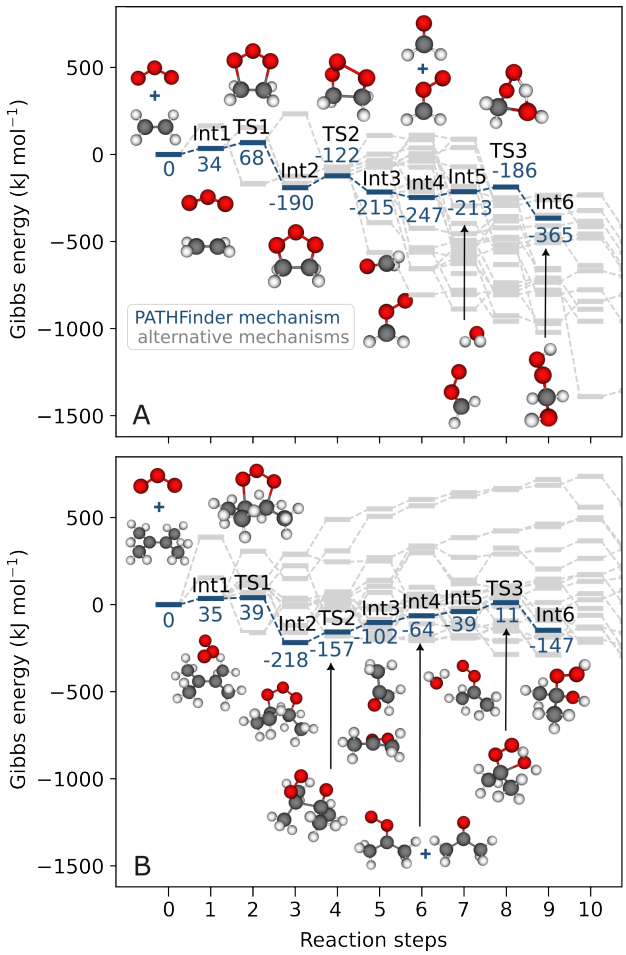


Figure 5: Reaction mechanism prediction by PATHFINDER: (A) from ozone and ethene to α-hydroxymethylhydroperoxide and from (B) ozone and tetramethylethene to 2-hydroxy-2-propylhydroperoxide. The mechanism with the lowest cost is depicted in dark blue color. Grey bars and lines represent alternative reaction mechanisms with higher compound costs.

droxyalkylhydroperoxides were defined arbitrarily as targets which will further decompose in water.[81] The selection of the same type of target for both CRNs allows for a direct comparison of the reaction mechanisms with experimental evidence (i.e., for reactions leading to compound **5** in Figure 1).

Figure 5A shows the most favorable mechanism from the reaction of ozone and ethene (source) to α-hydroxymethylhydroperoxide (target). The other 24 alternative mechanisms are depicted in grey because PATHFINDER classifies them as unlikely due to their greater compound costs (see elementary steps of reaction paths in Section S2.6.1). Mechanism 1 depicted in Figure 5A is spontaneous, with a total reaction Gibbs energy of –365 kJ·mol^{-1}.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

This reaction pathway starts with the barrierless formation of the vdW complex (Int1 in Figure 5). The subsequent 1,3-dipolar cycloaddition through TS1 exhibits an activation barrier of 34 kJ·mol^{-1}. As a result, the ozonide (Int2) is formed, with a favorable reaction energy of –258 kJ·mol^{-1}. Next, the ozonide decomposes through heterolytic O–O bond cleavage in TS2 with a barrier of +32 kJ·mol^{-1} to generate a flask composed by formaldehyde and methyliumperoxolate. After the dissociation of the flask, Int4, methyliumperoxolate reacts with water to generate a new flask, Int5. In the final reaction step, Int5 evolves through TS3, with an activation barrier of 27 kJ·mol^{-1}, to yield α-hydroxymethylhydroperoxide (Int6). Despite the fact that the PATHFINDER mechanism is not the lowest in energy (there are other pathways in Figure 5 with lower energy), it is indeed the shortest and the most favorable when considering both the activation barriers (i.e., kinetics) and initial concentrations of the reactants. Moreover, the alternative mechanisms show that there is a consistent trend in terms of total reaction Gibbs energy. For example, mechanism 25 (see Section S2.6.1) has a cost of 159.9 compared to 124.8 for mechanism 1. The difference stems from the fact that mechanism 25 includes an additional reaction step. For CRN-E, the formation of the α-hydroxymethylhydroperoxide is exergonic for all the reaction pathways shown in Figure 5A.

The most favorable reaction mechanism from ozone and tetramethylethene to 2-hydroxy-2-propylhydroperoxide is shown in Figure 5B. 24 alternative mechanisms are depicted in grey and exhibit greater compound costs. For instance, mechanism 25 illustrated in Section S2.6.2 has a cost of 288.5 as compared to 104.4 for mechanism 1. This large difference is not only caused by the extra reaction step of mechanism 25, but also by the formation of hydrogen which is not observed in experiments.[38] Mechanism 1 starts with the barrierless formation of the vdW complex (Int1) and then undergoes a 1,3-dipolar cycloaddition through TS1 with an activation energy of 4 kJ·mol^{-1} to generate the Criegee ozonide (Int2). Note that the activation barrier for TS1 from tetramethylethene is 30 kJ·mol^{-1} lower than TS1 from ethene, in good agreement with the higher reactivity of tetramethylethene reported in

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

experiments[38] (see $k_{O_3}^{exp}$-values in Figure 4). The ozonide decomposes through the heterolytic cleavage reaction, via TS2, to form a flask composed by acetone and 2-peroxopropylate. The flask dissociates to form Int4. Subsequently, 2-peroxopropylate forms yet a new flask with water, Int5 which yields 2-hydroxy-2-propylhydroperoxide through TS3 with an activation barrier of 28 kJ·mol^{-1}. This reaction mechanism is not spontaneous because the ozonide (Int2) is 71 kJ·mol^{-1} more stable than the final product 2-hydroxy-2-propylhydroperoxide. Alternative reaction pathways found by PATHFINDER also show the same endergonic trend towards the formation of the 2-hydroxy-2-propylhydroperoxide. This observation reveals the principle difference between the reaction mechanisms of ozone with tetramethylethene and ethene, respectively. While the formation of hydroxyalkylhydroperoxides is favored for CRN-E, it is not for CRN-T. This interpretation perfectly agrees with the experimental data, as α-hydroxymethylhydroperoxide is one of the main products whereas 2-hydroxy-2-propylhydroperoxide is not.[38]

Microkinetic Analyses

While the reaction pathway identification relies on an explicit definition of source and target compounds, microkinetic simulations are literally exploratory within the network of generated compounds, flasks, and transition states. Here, we employed KINETX[34] to carry out microkinetic analyses for CRN-E and CRN-T and the outcome is depicted in Figure 6. The initial conditions are identical to those used for reaction pathway identification from experiments of Dowideit and von Sonntag[38] with excess of olefins and water. The kinetic parameters and initial concentrations are compiled in Table S2.

Figure 6A shows the reaction kinetics of compounds exceeding femto-molar concentrations in CRN-E that included all compounds and reactions included in Table 1. Typical intermediates of ethene ozonolysis (i.e., the ozonide, α-hydroxymethyl hydroperoxide, methylium peroxolate) are formed continuously until they decrease as ozone is consumed completely (between 1 to 100 seconds). Formaldehyde and α-hydroxymethyl hydroperoxide are the fi-

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 ORCID: https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

nal products of this reaction and they form in the correct equimolar yields. This accurate reproduction of experimental transformation products[38] was obtained even though reaction rate constants, such as those for the ozone consumption ($k_{O_3}^{QC}$ in Figure 4), deviate orders of magnitude from measured rate constants. Our data thus confirms the error cancellation in transition state energy calculations invoked above.

Note that we combined DLPNO-CCSD(T) electronic energies with LC-PBE entropy corrections in our microkinetic simulations to calculate Gibbs energies.[82,83] When reaction rates are exclusively calculated with LC-PBE Gibbs energies, the resulting kinetic simulations lead to the formation of the Criegee ozonide as predominant final product (Figures S11A and S12A for CRN-E and CRN-T, respectively). None of these calculations match experimental evidence.[38] These observations reinforces results from our QC method benchmarking (Figure 2B) which show that LC-PBE electronic energies do not ensure chemical accuracy in simulations of ozone chemistry. When reaction rate constants are calculated us-

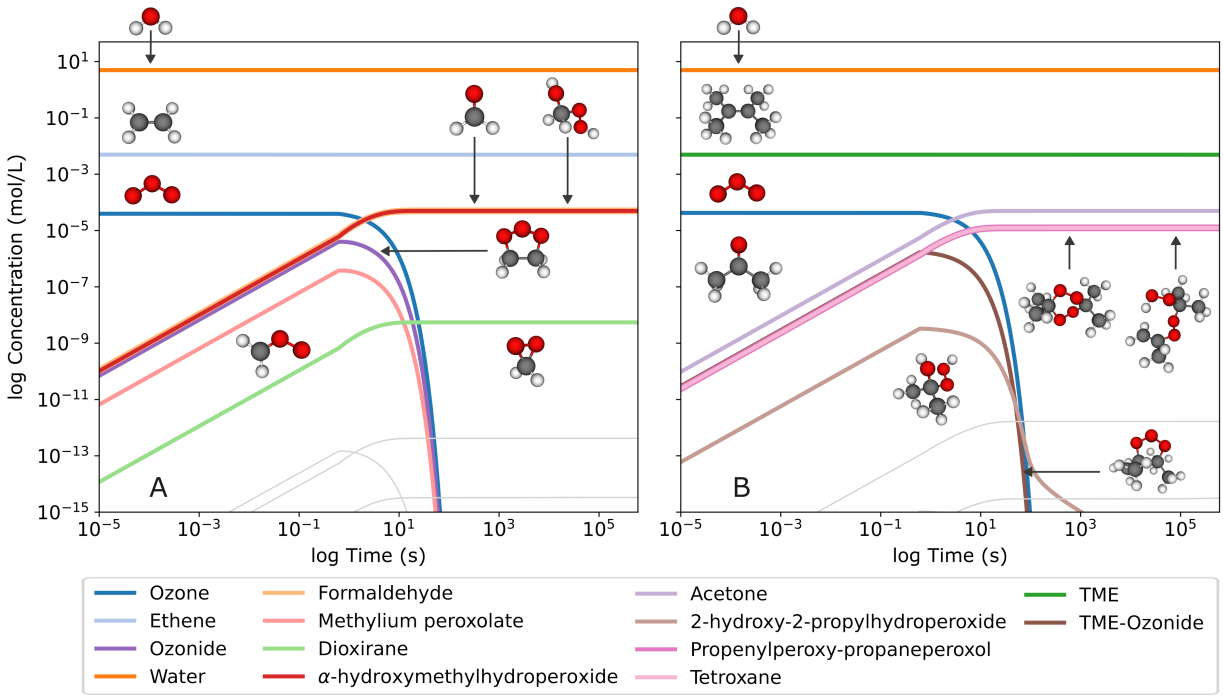


Figure 6: Concentration–time plots derived from the microkinetic simulations of (A) CRN-E and (B) CRN-T. Only compounds at concentrations exceeding concentrations of 10^{-10} M at any time of the simulation are shown in color lines.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

ing DLPNO-CCSD(T) electronic energies only, the microkinetic simulations form the correct compounds but in flasks instead of compounds (Figures S11B and S11B). Electronic energies do not favor dissociation reactions of flasks due to the lack of the entropic contribution.

Figure 6B shows the kinetics of compounds in CRN-T from a microkinetic simulation carried out in analogous fashion as for CRN-E (see Table 1 for number of compounds and reactions and Table S2 for simulation parameters). This microkinetic analysis is based on exploration data with a confined set of identified compounds (see above). Even though the lower depth of CRN-T is a limitation to the network's evaluation, these results allow for critical insights into the interpretation of automatic CRN explorations. In CRN-T, tetramethylethene and ozone react to acetone as primary product through decomposition of the tetramethylethene-ozonide via heterolytic O–O bond cleavage in agreement with experiments.[38] The fact that the hydroxy-alkyl hydroperoxide (2-hydroxypropyl-2-hydroperoxide) does not become the predominant product is also consistent with experiments. However, formation of two unexpected compounds, 3,3,6,6-tetramethyl-1,2,4,5-tetroxane, simply referred to as tetroxane (Gibbs energy of –227 kJ·mol^{-1}, Figure S13A) and propenylperoxypropaneperoxol (–277 kJ·mol^{-1}, Figure S13B) vs. –218 kJ·mol^{-1} for the ozonide (Figure 5) hint at limited extent of reaction progress or lacking depth of CRN-T exploration. We hypothesize that this outcome is also due to an artifact with regard to the role that water plays in the hydrolysis of the alkyliumperoxolate (i.e., the 2-peroxo propylate zwitterion in CRN-T) to hydrogen peroxide and acetone (**3** in Figure 1).

As discussed in Section S2.11, the transition state leading to the formation of hydrogen peroxide from 2-peroxopropylate strongly depends on the position of the attacking water (Figure S15). The need for aligning water to the hydrolysis reaction coordinate artificially increases the activation energy thus ultimately hindering the reaction. We note that this issue does not systematically affect reaction of water in general. For example, in the hydrolysis reaction with 2-peroxopropylate to form 2-hydroxy-2-propylhydroperoxide, water attacks from an optimal angle, leading to a smooth trajectory on the potential energy sur-

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

face (Figure S17). Our observations point to intrinsic limitations of the implicit solvation model (CPCM) to describe local solvation effects.[84] Developments for the implementation of updated solvation methodology[85,86] into CHEMOTON are currently underway.[87]

Significance

Our study demonstrates how quantum chemistry-based CRN exploration offers novel avenues to elucidate reactions of ozone with organic compounds in aqueous solution. Insights from the evaluation of reactions of tetramethylethene with ozone, however, revealed bottlenecks that need to be addressed in improved computational and data processing workflows for more expansive CRN exploration of chemical oxidation processes. Examples include the efficient combination of fast semiempirical electronic structure calculations with DFT and CC methods to arrive at reliable structures and minimum/TS energies of structurally more complex compounds than substituted ethenes, the management of the ensuing large datasets for evaluation and visualization, as well as the handling of water as both solvent and reactant.

Once established, CRN evaluations of chemical oxidation processes can support the re-evaluation of reactions of many organic compounds of concern in oxidative water and wastewater treatment processes. Even though sophisticated chemical analyses, such as non-target analysis by high resolution mass spectrometry, generates numerous structural features of organic transformation products,[88] such approaches are unable to provide a comprehensive view of the identity of the numerous intermediates and products (e.g., from drinking water ozonation processes[15]). The numerous compounds acquired in CRN databases might prove highly beneficial for prediction of high resolution mass spectra, for example if combined with deep learning-based approaches developed for this purpose.[89]

Moreover, (eco)toxicity evaluations recommend circumventing the assignment of the observed toxicity to chemicals in polluted and treated waters because only a minor fraction of the observed toxicity can be allocated to individual compounds in mixtures of pollutants and their transformation products.[90] The prediction of transformation products and

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 ORCID: https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. License: CC BY 4.0

coupling this information to biodegradability assessment is a critial step in the evaluation of their abatement in biofilters, which are common after ozonation. Chemical reaction networks generated through high-throughput computational chemistry have great potential to address knowledge gaps in (eco)toxicity evaluations by providing atomistic evidence on transient chemical species and potentially overlooked transformation products. Finally, we envision that predictions of potentially harmful or toxic compounds or compound (sub)structures from CRN explorations of oxidative water treatment could become pivotal for the evaluation of environmental impacts of chemicals in early design phases.[91]

Data Availability Statement

Reaction network data is available at Zenodo,[92] and can be downloaded as (i) MongoDB files to then import them using SCINE, and (ii) self-contained HTML files for interactive data visualization directly in a browser. In addition, molecular structures of the CRNs are accessible at ioChem-BD,[93] via the associated collection.[94]

Supporting Information Available

Short introduction to automated chemical reaction network explorations, additional results for computational method development, reaction network exploration outcomes, reaction pathways identification, and microkinetic modeling.

Acknowledgments

We are grateful to Jan P. Unsleber and Thomas Weymuth for introducing SCINE to the first author at the start of the project. We also thank Miguel Steiner and Paul L. Türtscher for their helpful discussions on the use of Chemoton.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

References

(1) von Gunten, U. Oxidation processes in water treatment: Are We on track? *Environ. Sci. Technol.* **2018**, *52*, 5062–5075, https://doi.org/10.1021/acs.est.8b00586.

(2) Von Gunten, U. Oxidation processes and me. *Water Research* **2024**, *253*, 121148, https://doi.org/10.1016/j.watres.2024.121148.

(3) Li, X.-F.; Mitch, W. A. Drinking water disinfection byproducts (DBPs) and human health effects: Multidisciplinary challenges and opportunities. *Environ. Sci. Technol.* **2018**, *52*, 1681–1689, https://doi.org/10.1021/acs.est.7b05440.

(4) Sedlak, D. L.; von Gunten, U. The chlorine dilemma. *Science* **2011**, *331*, 42–43, https://doi.org/10.1126/science.119639.

(5) Lee, Y.; Sedlak, D. L.; von Gunten, U. Oxidative water treatment: the track ahead. *Environ. Sci. Technol.* **2023**, *57*, 18391–18392, https://doi.org/10.1021/acs.est.3c07785.

(6) Lim, S.; Shi, J. L.; von Gunten, U.; McCurry, D. L. Ozonation of organic compounds in water and wastewater: A critical review. *Water Res.* **2022**, *213*, 118053, https://doi.org/10.1016/j.watres.2022.118053.

(7) Prasse, C.; Ford, B.; Nomura, D. K.; Sedlak, D. L. Unexpected transformation of dissolved phenols to toxic dicarbonyls by hydroxyl radicals and UV light. *Proceedings of the National Academy of Sciences of the United States of America* **2018**, *115*, 2311–2316, https://doi.org/10.1073/pnas.1715821115.

(8) Prasse, C.; von Gunten, U.; Sedlak, D. L. Chlorination of phenols revisited: Unexpected formation of α,β-unsaturated C_4-dicarbonyl ring cleavage products. *Environ. Sci. Technol.* **2020**, *54*, 826–834, https://doi.org/10.1021/acs.est.9b04926.

(9) Mitch, W. A.; Richardson, S. D.; Zhang, X.; Gonsior, M. High-molecular-

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

weight by-products of chlorine disinfection. *Nature Water* **2023**, *1*, 336–347, https://doi.org/10.1038/s44221-023-00064-x.

(10) Marron, E. L.; Prasse, C.; Buren, J. V.; Sedlak, D. L. Formation and fate of carbonyls in potable water reuse systems. *Environ. Sci. Technol.* **2020**, *54*, 10895–10903, https://doi.org/10.1021/acs.est.0c02793.

(11) Houska, J.; Manasfi, T.; Gebhardt, I.; von Gunten, U. Ozonation of lake water and wastewater: Identification of carbonous and nitrogenous carbonyl-containing oxidation byproducts by non-target screening. *Wat. Res.* **2023**, *232*, 119484, https://doi.org/10.1016/j.watres.2022.119484.

(12) Manasfi, T.; Houska, J.; Gebhardt, I.; von Gunten, U. Formation of carbonyl compounds during ozonation of lake water and wastewater: Development of a non-target screening method and quantification of target compounds. *Wat. Res.* **2023**, *237*, 119751, https://doi.org/10.1016/j.watres.2023.119751.

(13) Marron, E. L.; Mitch, W. A.; von Gunten, U.; Sedlak, D. L. A tale of two treatments: The multiple barrier approach to removing chemical contaminants during potable water reuse. *Acc. Chem. Res.* **2019**, *52*, 615–622, https://doi.org/10.1021/acs.accounts.8b00612.

(14) Formation of transformation products during ozonation of secondary wastewater effluent and their fate in post-treatment: From laboratory- to full-scale. *Water Res.* **2021**, *200*, 117200, https://doi.org/10.1016/j.watres.2021.117200.

(15) Gulde, R.; Clerc, B.; Rutsch, M.; Helbing, J.; Salhi, E.; McArdell, C. S.; von Gunten, U. Oxidation of 51 micropollutants during drinking water ozonation: Formation of transformation products and their fate during biological post-filtration. *Water Res.* **2021**, *207*, 117812, https://doi.org/10.1016/j.watres.2021.117812.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(16) Ferraro, P. J.; Prasse, C. Reimagining safe drinking water on the basis of twenty-first-century science. *Nat Sustain* **2021**, *4*, 1032–1037, https://doi.org/10.1038/s41893-021-00760-0.

(17) Rosario-Ortiz, F.; Rose, J.; Speight, V.; von Gunten, U.; Schnoor, J. How do you like your tap water? *Science* **2016**, *351*, 912–914, https://doi.org/10.1126/science.aaf0953.

(18) Lee, M.; Blum, L. C.; Schmid, E.; Fenner, K.; von Gunten, U. A computer-based prediction platform for the reaction of ozone with organic compounds in aqueous solution: kinetics and mechanisms. *Environ. Sci.: Processes Impacts* **2017**, *19*, 465–476, https://doi.org/10.1039/C6EM00584E.

(19) Kamath, D.; Mezyk, S. P.; Minakata, D. Elucidating the elementary reaction pathways and kinetics of hydroxyl radical-induced acetone degradation in aqueous phase advanced oxidation processes. *Environ. Science Technol.* **2018**, *52*, 7763–7774, https://doi.org/10.1021/acs.est.8b00582.

(20) von Sonntag, C.; von Gunten, U. *Chemistry of ozone in water and wastewater treatment: from basic principles to applications*; IWA Publishing, 2012; https://doi.org/10.2166/9781780400839.

(21) Guo, X.; Minakata, D.; Niu, J.; Crittenden, J. Computer-based first-principles kinetic modeling of degradation pathways and byproduct fates in aqueous-phase advanced oxidation processes. *Environ. Sci. Technol.* **2014**, *48*, 5718–5725, https://doi.org/10.1021/es500359g.

(22) Tentscher, P. R.; Lee, M.; von Gunten, U. Micropollutant oxidation studied by quantum chemical computations: methodology and applications to thermodynamics, kinetics, and reaction mechanisms. *Acc. Chem. Res.* **2019**, *52*, 605–614, https://doi.org/10.1021/acs.accounts.8b00610.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(23) Trogolo, D.; Arey, J. S.; Tentscher, P. R. Gas-phase ozone reactions with a structurally diverse set of molecules: barrier heights and reaction energies evaluated by coupled cluster and density functional theory calculations. *J. Phys. Chem. A* **2019**, *123*, 517–536, https://doi.org/10.1021/acs.jpca.8b10323.

(24) Trogolo, D.; Mishra, B. K.; Heeb, M. B.; von Gunten, U.; Arey, J. S. Molecular mechanism of NDMA formation from *N,N*-dimethylsulfamide during ozonation: quantum chemical insights into a bromide-catalyzed pathway. *Environ. Sci. Technol.* **2015**, *49*, 4163–4175, https://doi.org/10.1021/es504407h.

(25) Minakata, D.; von Gunten, U. Predicting transformation products during aqueous oxidation processes: current state and outlook. *Environ. Sci. Technol.* **2023**, *57*, 18410–18419, https://doi.org/10.1021/acs.est.3c04086.

(26) Weymuth, T.; Unsleber, J. P.; Türtscher, P. L.; Steiner, M.; Sobez, J.-G.; Müller, C. H.; Mörchen, M.; Klasovita, V.; Grimmel, S. A.; Eckhoff, M.; Csizi, K.-S.; Bosia, F.; Bensberg, M.; Reiher, M. SCINE—Software for chemical interaction networks. *J. Chem. Phys.* **2024**, *160*, 222501, https://doi.org/10.1063/5.0206974.

(27) Unsleber, J. P.; Liu, H.; Talirz, L.; Weymuth, T.; Mörchen, M.; Grofe, A.; Wecker, D.; Stein, C. J.; Panyala, A.; Peng, B.; Kowalski, K.; Troyer, M.; Reiher, M. High-throughput ab initio reaction mechanism exploration in the cloud with automated multi-reference validation. *J. Chem. Phys.* **2023**, *158*, 084803, https://doi.org/10.1063/5.0136526.

(28) Baiardi, A.; Grimmel, S. A.; Steiner, M.; Türtscher, P. L.; Unsleber, J. P.; Weymuth, T.; Reiher, M. Expansive quantum mechanical exploration of chemical reaction paths. *Acc. Chem. Res.* **2022**, *55*, 35–43, https://doi.org/10.1021/acs.accounts.1c00472.

(29) Unsleber, J. P.; Reiher, M. The exploration of chemical reaction networks. *Annu. Rev.*

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Phys. Chem. **2020**, *71*, 121–142, ,https://doi.org/10.1146/annurev-physchem-071119-040123.

(30) Unsleber, J. P.; Grimmel, S. A.; Reiher, M. Chemoton 2.0: Autonomous exploration of chemical reaction networks. *J. Chem. Theory Comput.* **2022**, *18*, 5393–5409, https://doi.org/10.1021/acs.jctc.2c00193.

(31) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB - An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671, https://doi.org/10.1021/acs.jctc.8b01176.

(32) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIRES Comp. Mol. Sci.* **2021**, *11*, e1493, https://doi.org/10.1002/wcms.1493.

(33) Türtscher, P. L.; Reiher, M. Pathfinder - Navigating and analyzing chemical reaction networks with an efficient graph-based approach. *J. Chem. Inf. Model.* **2023**, *63*, 147–160, https://doi.org/10.1021/acs.jcim.2c01136.

(34) Proppe, J.; Reiher, M. Mechanism deduction from noisy chemical reaction networks. *J. Chem. Inf. Model.* **2019**, *15*, 357–370, https://doi.org/10.1021/acs.jctc.8b00310.

(35) Steiner, M.; Reiher, M. A human-machine interface for automatic exploration of chemical reaction networks. *Nat. Comm.* **2024**, *15*, 3680, https://doi.org/10.1038/s41467-024-47997-9.

(36) Bensberg, M.; Reiher, M. Concentration-flux-steered mechanism exploration with an organocatalysis application. *Israel Journal of Chemistry* **2023**, *63*, e202200123, 10.1002/ijch.202200123.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(37) Csizi, K.-S.; Steiner, M.; Reiher, M. Nanoscale chemical reaction exploration with a quantum magnifying glass. *Nat. Comm.* **2024**, *15*, 5320, https://doi.org/10.1038/s41467-024-49594-2.

(38) Dowideit, P.; von Sonntag, C. Reaction of ozone with ethene and its methyl- and chlorine-substituted derivatives in aqueous solution. *Environ. Sci. Technol.* **1998**, *32*, 1112–1119, https://doi.org/10.1021/es971044j.

(39) Bailey, P. S. The reactions of ozone with organic compounds. *Chem. Rev.* **1958**, *58*, 925–1010, https://doi.org/10.1021/cr50023a005.

(40) Miliordos, E.; Xantheas, S. S. On the bonding nature of ozone (O_3) and Its sulfur-substituted analogues SO_2, OS_2, and S_3: Correlation between their biradical character and molecular properties. *J. Am. Chem. Soc.* **2014**, *136*, 2808–2817, https://doi.org/10.1021/ja410726u.

(41) Wheeler, S. E.; Ess, D. H.; Houk, K. N. Thinking out of the black box: accurate barrier heights of 1,3-dipolar cycloadditions of ozone with acetylene and ethylene. *J. Phys. Chem. A* **2008**, *112*, 1798–1807, https://doi.org/10.1021/jp710104d.

(42) Zhao, Y.; Tishchenko, O.; Gour, J. R.; Li, W.; Lutz, J. J.; Piecuch, P.; Truhlar, D. G. Thermochemical kinetics for multireference systems: addition reactions of ozone. *J. Phys. Chem. A* **2009**, *113*, 5786–5799, https://doi.org/10.1021/jp811054n.

(43) Neese, F. Software update: the ORCA program system—Version 5.0. *WIREs Comput Mol Sci.* **2022**, *12*, e1606, https://doi.org/10.1002/wcms.1606.

(44) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948, https://doi.org/10.1021/ct100684s.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(45) Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **2007**, *13*, 1173–1213, https://doi.org/10.1007/s00894-007-0233-4.

(46) Bensberg, M.; Reiher, M. Corresponding active orbital spaces along chemical reaction paths. *J. Phys. Chem. Lett.* **2023**, *14*, 2112–2118, https://doi.org/10.1021/acs.jpclett.2c03905.

(47) Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. *Chem. Rev.* **2012**, *112*, 108–181, https://doi.org/10.1021/cr200137a.

(48) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822–8824, https://doi.org/10.1103/PhysRevB.33.8822.

(49) Perdew, J. P. Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *34*, 7406–7406, https://doi.org/10.1103/PhysRevB.34.7406.

(50) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170, https://doi.org/10.1063/1.478522.

(51) Vydrov, O. A.; Scuseria, G. E. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.* **2006**, *125*, 234109, https://doi.org/10.1063/1.2409292.

(52) Schäfer, A.; Huber, C.; Ahlrichs, R. Fully optimized contracted gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* **1994**, *100*, 5829–5835, https://doi.org/10.1063/1.467146.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(53) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104, https://doi.org/10.1063/1.3382344.

(54) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J Comput Chem* **2011**, *32*, 1456–1465, https://doi.org/10.1002/jcc.21759.

(55) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483, https://doi.org/10.1016/S0009-2614(89)87395-6.

(56) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139*, 134101, https://doi.org/10.1063/1.4821834.

(57) Dunning, J., Thom H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023, https://doi.org/10.1063/1.456153.

(58) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J Comput Chem* **2003**, *24*, 669–681, https://doi.org/10.1002/jcc.10189.

(59) Besora, M.; Maseras, F. Microkinetic modeling in homogeneous catalysis. *WIRES Comp. Mol. Sci.* **2018**, *8*, 1–13, https://doi.org/10.1002/wcms.1372.

(60) Bensberg, M.; Grimmel, S. A.; Lang, L.; Simm, G. N.; Sobez, J.-G.; Steiner, M.; Türtscher, P.; Unsleber, J. P.; Weymuth, T.; Reiher, M. qcscine/chemoton: Release 3.0.0. 2023-05-12; https://doi.org/10.3929/ethz-b-000611922.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(61) Maeda, S.; Harabuchi, Y.; Takagi, M.; Taketsugu, T.; Morokuma, K. Artificial force induced reaction (AFIR) method for exploring quantum chemical potential energy surfaces. *The Chemical Record* **2016**, *16*, 2232–2248, https://doi.org/10.1002/tcr.201600043.

(62) Quapp, W.; Bofill, J. M. Some mathematical reasoning on the artificial force induced reaction method. *J. Comput. Chem* **2020**, *41*, 629–634, https://doi.org/10.1002/jcc.26115.

(63) Bensberg, M.; Brunken, C.; Csizi, K.-S.; Grimmel, S. A.; Gugler, S.; Sobez, J.-G.; Steiner, M.; Türtscher, P. L.; Unsleber, J. P.; Weymuth, T.; Reiher, M. qcscine/puffin: Release 1.2.0. 2023; `https://doi.org/10.5281/zenodo.7928099`, https://doi.org/10.5281/zenodo.7928099.

(64) Bensberg, M.; Grimmel, S. A.; Sobez, J.-G.; Steiner, M.; Unsleber, J. P.; Reiher, M. qcscine/database: Release 1.1.0. 2022-08-15; https://doi.org/10.3929/ethz-b-000563724.

(65) Bensberg, M.; Brunken, C.; Csizi, K.-S.; Grimmel, S.; Gugler, S.; Sobez, J.-G.; Steiner, M.; Türtscher, P. L.; Unsleber, J. P.; Vaucher, A. C.; Weymuth, T.; Reiher, M. qcscine/readuct: Release 5.1.0. 2023; `https://doi.org/10.5281/zenodo.10159561`, https://doi.org/10.5281/zenodo.10159561.

(66) Bensberg, M.; Grimmel, S. A.; Sobez, J.-G.; Steiner, M.; Unsleber, J. P.; Reiher, M. qcscine/molassembler: Release 2.0.0. 2023; `https://doi.org/10.5281/zenodo.7928074`, https://doi.org/10.5281/zenodo.7928074.

(67) Baiardi, A. et al. qcscine/utilities: Release 8.0.0. 2023; `https://doi.org/10.5281/zenodo.7928050`, https://doi.org/10.5281/zenodo.7928050.

(68) Müller, C. H.; Steiner, M.; Unsleber, J. P.; Weymuth, T.; Bensberg, M.; Csizi, K.-S.; Mörchen, M.; Türtscher, P. L.; Reiher, M. Heron: Visualizing and Controlling

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Chemical Reaction Explorations and Networks. *J. Phys. Chem. A* **2024**, *128*, 9028–9044, https://doi.org/10.1021/acs.jpca.4c03936.

(69) Bensberg, M.; Brandino, G. P.; Can, Y.; Del, M.; Grimmel, S. A.; Mesiti, M.; Müller, C. H.; Steiner, M.; Türtscher, P. L.; Unsleber, J. P.; Weberndorfer, M.; Weymuth, T.; Reiher, M. qcscine/heron: Release 1.0.0. 2022; `https://doi.org/10.5281/zenodo.7038388`.

(70) Petrus, E.; Garay-Ruiz, D.; Bo, C.; Reiher, M.; Hofstetter, T. VizChemoton v1.0.0. 2025; `https://doi.org/10.5281/zenodo.14803803`.

(71) Criegee, R. Mechanism of ozonolysis. *Angew. Chem. Int. Ed.* **1975**, *14*, 745–752, https://doi.org/10.1002/anie.197507451.

(72) Friedrich, N.-O.; de Bruyn Kops, C.; Flachsenberg, F.; Sommer, K.; Rarey, M.; Kirchmair, J. Benchmarking commercial conformer ensemble generators. *J. Chem. Inf. Model.* **2017**, *57*, 2719–2728, https://doi.org/10.1021/acs.jcim.7b00505.

(73) Jakobsen, S.; Kristensen, K.; Jensen, F. Electrostatic potential of insulin: exploring the limitations of density functional theory and force field methods. *J. Chem. Theory Comput.* **2013**, *9*, 3978–3985, https://doi.org/10.1021/ct400452f.

(74) Braïda, B.; Galembeck, S. E.; Hiberty, P. C. Ozone and Other 1,3-Dipoles: Toward a Quantitative Measure of Diradical Character. *J. Chem. Theory Comput. 13*, 3228–3235.

(75) Řezáč, J.; Fanfrlík, J.; Salahub, D.; Hobza, P. Semiempirical quantum chemical PM6 method augmented by dispersion and H-bonding correction terms reliably describes various types of noncovalent complexes. *J. Chem. Theory Comput.* **2009**, *5*, 1749–1760, https://doi.org/10.1021/ct9000922.

(76) Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approxima-

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

tions for the exchange-correlation energy. *AIP Conf. Proc* **2001**, *577*, 1–20, https://doi.org/10.1063/1.1390175.

(77) Pople, J. A. Nobel Lecture: Quantum chemical models. *Rev. Mod. Phys.* **1999**, *71*, 1267–1274, https://doi.org/10.1103/RevModPhys.71.1267.

(78) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101, https://doi.org/10.1063/1.5011798.

(79) Proppe, J.; Kircher, J. Uncertainty Quantification of Reactivity Scales**. *ChemPhysChem* **2022**, *23*, e202200061, https://doi.org/10.1002/cphc.202200061.

(80) Ariai, J.; Gellrich, U. The entropic penalty for associative reactions and their physical treatment during routine computations. *Phys. Chem. Chem. Phys.* **2023**, *25*, 14005–14015, https://doi.org/10.1039/D3CP00970J.

(81) Qiu, J.; Tonokura, K.; Enami, S. Proton-catalyzed decomposition of α-Hydroxyalkyl-hydroperoxides in water. *J. Chem. Theory Comput.* **2020**, *54*, 10561–10569, https://doi.org/10.1021/acs.est.0c03438.

(82) Kurfman, L. A.; Odbadrakh, T. T.; Shields, G. C. Calculating Reliable Gibbs Free Energies for Formation of Gas-Phase Clusters that Are Critical for Atmospheric Chemistry: (H2SO4)3. *The Journal of Physical Chemistry A* **2021**, *125*, 3169–3176, PMID: 33825467.

(83) Afify, N. D.; Sweatman, M. B. Solvent-mediated modification of thermodynamics and kinetics of monoethanolamine regeneration reaction in amine-stripping carbon capture: Computational chemistry study. *J. Chem. Phys.* **2024**, *160*, 014501, https://doi.org/10.1063/5.0169382.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

(84) Basdogan, Y.; Maldonado, A. M.; Keith, J. A. Advances and challenges in modeling solvated reaction mechanisms for renewable fuels and chemicals. *WIREs Computational Molecular Science* **2020**, *10*, e1446.

(85) Simm, G. N.; Türtscher, P. L.; Reiher, M. Systematic microsolvation approach with a cluster-continuum scheme and conformational sampling. *J. Comp. Chem.* **2020**, *41*, 1144–1155, https://doi.org/10.1002/jcc.26161.

(86) Bensberg, M.; Türtscher, P. L.; Unsleber, J. P.; Reiher, M.; Neugebauer, J. Solvation free energies in subsystem density functional theory. *J. Chem. Theory Comput.* **2022**, *18*, 723–740, https://doi.org/10.1021/acs.jctc.1c00864.

(87) Türtscher, P. L.; Reiher, M. Automated Microsolvation for Minimum Energy Path Construction in Solution. 2025; `https://arxiv.org/abs/2502.07965`.

(88) Hollender, J. et al. NORMAN guidance on suspect and non-target screening in environmental monitoring. *Environ Sci. Eur.* **2023**, *35*, 75, https://doi.org/10.1186/s12302-023-00779-4.

(89) Overstreet, R.; King, E.; Clopton, G.; Nguyen, J.; Ciesielski, D. QC-GN2oMS2: a graph neural net for high resolution mass spectra prediction. *J. Chem. Inf. Model.* **2024**, *64*, 5806–5816, https://doi.org/10.1021/acs.jcim.4c00446.

(90) Escher, B. I. et al. Modernizing persistence–bioaccumulation–toxicity (PBT) assessment with high throughput animal-free methods. *Arch. Toxicol.* **2023**, *97*, 1267–1283, https://doi.org/10.1007/s00204-023-03485-5.

(91) Zimmerman, J. B.; Anastas, P. T.; Erythropel, H. C.; Leitner, W. Designing for a green chemistry future. *Science* **2020**, *367*, 397–400, https://doi.org/10.1126/science.aay3060.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

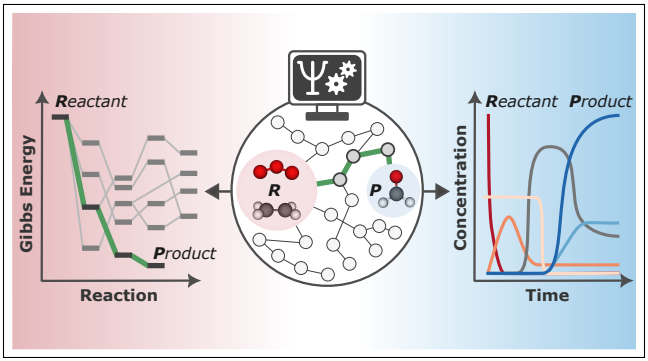
(92) Petrus, E.; Hunkeler, L. A.; von Gunten, U.; Reiher, M.; Hofstetter, T. B. Ozonation exploration data set: ethene and tetramethylethene in water. 2024; https://zenodo.org/records/14448899.

(93) Álvarez-Moreno, M.; De Graaf, C.; López, N.; Maseras, F.; Poblet, J. M.; Bo, C. Managing the computational chemistry big data problem: The ioChem-BD platform. *J. Chem. Inf. Model.* **2015**, *55*, 95–103, https://doi.org/10.1021/ci500593j.

(94) Petrus, E. Automated reaction exploration of ozonation processes for model olefins in water. 2024; http://dx.doi.org/10.19061/iochem-bd-6-430, Accessed: 2024-12-13.

https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

TOC Graphic



https://doi.org/10.26434/chemrxiv-2025-cf5pg-v2 **ORCID:** https://orcid.org/0000-0003-1906-367X Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0