More and Faster: Simultaneously Improving Reaction Coverage and Computational Cost in Automated Reaction Prediction Tasks

Qiyuan Zhao and Brett M. Savoie*

Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, 47906

E-mail: bsavoie@purdue.edu

Abstract

Automated reaction prediction has the potential to elucidate complex reaction networks for applications ranging from combustion to materials degradation. Although substantial progress has been made in predicting specific reaction pathways and resolving mechanisms, the computational cost and inconsistent reaction coverage of automated prediction are still obstacles to exploring deep reaction networks without using heuristics. Here we show that cost can be reduced and reaction coverage can be increased simultaneously by relatively straightforward modifications of the reaction enumeration, geometry initialization, and transition state convergence algorithms that are common to many emerging prediction methodologies. These changes are implemented in the context of Yet Another Reaction Program (YARP), our reaction prediction package, for which we report a head-to-head comparison with prevailing methods for two benchmark reaction prediction tasks. In all cases, we observe near perfect recapitulation of established reaction pathways and products by YARP, without the use of heuristics or other domain knowledge to guide reaction selection. In addition, YARP also discovers many new kinetically relevant pathways and products reported here for the first time. This is achieved while simultaneously reducing the cost of reaction characterization nearly 100-fold and increasing transition state success rates and intended rates over 2-fold and 10-fold, respectively, compared with recent benchmarks. This combination of ultra-low cost and high reaction-coverage creates opportunities to explore the reactivity of larger systems and more complex reaction networks for applications like chemical degradation, where approaches based on domain heuristics fail.

1 Introduction

The reaction network prediction problem consists of predicting the transition states, kinetically relevant intermediates, and products for a set of reactants. Decades of research has been devoted to this topic for specific applications, ranging from the evaluation of combustion pathways,[1–5] cellular metabolism,[6–10] and atmospheric chemistry,[11–15] to the related inverse problem of retrosynthetic organic reaction planning (i.e., generating a reaction network in reverse).[16–20] Although the details are specific to each application, the problem common to all is resolving which reactions happen and when as a function of relevant environmental variables (i.e., temperature, pressure, concentrations, reagents, phase, etc). For applications where sufficient domain knowledge of plausible reactions exists, workable solutions have been developed to algorithmically generate reaction networks that are then refined with feedback from experiments or additional computational characterizations.[21–23] More recently, machine learning has also enabled major advances in data-rich reaction problems, with demonstrations of models capable of predicting retrosynthetic pathways for complex organic molecules that are competitive with the best expert systems and chemical intuition.[24,25] However, for problems where an established set of reactions and reaction data do not exist, a fundamentally different approach is required for reaction network prediction.

In recent years, several groups have recognized the opportunity to automatically elucidate reaction networks absent previous domain knowledge by leveraging now mature quantum-chemistry based transition state characterizations[26–32] and modern computational throughput. As has been recently reviewed in several places,[22,33–35] these approaches can roughly be categorized on the basis of whether they utilize single-ended transition state searches (e.g., ADNN,[36] AFIR,[37,38] SSWM[39,40]), double-ended transition state searches (e.g., ZStruct[41,42] and the methods of Suleimanov and Green[43,44]), or reaction templates (e.g., NetGen,[45] RMG,[46] KinBot[47]) to drive reaction discovery, with each coupled to a suitably accurate model chemistry to locate transition states and establish the thermodynamics of products. Although still in a relatively early phase, there are already many demonstrations of such algorithms automatically recapitulating established reaction mechanisms in benchmark systems, as well as authentic predictions of reaction pathways that were previously undocumented.[48–56] Despite this success, computational cost still represents a serious impediment to characterizing complex reaction networks involving large numbers of atoms or reactions occurring in condensed phases. In particular, the underlying reaction exploration algorithm must be sufficiently general such that all kinetically relevant pathways are identified and characterized by quantum chemistry. As such, reaction exploration cannot be entirely naive, as the range of possible bond rearrangements grows roughly factorially with the number of atoms in the reactants and would thus overwhelm even the most economical model chemistry.[23] On the other hand, it has been observed that the number of kinetically relevant reactions grows linearly with respect to number of intermediates for established reaction networks across many domains.[57,58] This implies that the kinetically relevant reactions in a typical network are computationally tractable to characterize, although at present there is no *a priori* method for robustly distinguishing physical from unphysical reactions besides subjecting them to costly quantum chemistry calculations or relying on heuristics based on domain knowledge (i.e., the very thing that is missing in many motivating applications for computational prediction).

In the current work, we introduce several strategies for improving the accuracy and size of reaction networks that can be computationally characterized without using domain heuristics. First, we employ the recently developed semi-empirical GFN2-xTB model chemistry of Grimme[59,60] to pre-optimize reaction pathways prior to more expensive density functional theory (DFT) characterizations. Although GFN2-xTB predicts inaccurate transition state energies, we observe that pre-optimization with GFN2-xTB yields excellent initial guesses for reaction pathways that reduces the number of required DFT gradient calls nearly 100-fold compared with recent DFT benchmarks. Second, we implement an elementary reaction step that is generally applicable to neutral closed-shell organic chemistry. We demonstrate that this elementary step exhibits more comprehensive reaction discovery and lower activation energy pathways than other recent graph-based enumeration schemes. Third, we establish a robust initial structure optimization procedure based on jointly optimizing the reactant and product geometries prior to the transition state search. Together, these two relatively straightforward strategies lead to dramatic improvements in the success rate and intended rates of transition state convergence. These features are implemented in the context of a recursive reaction enumeration and double-ended transition state search algorithm, which we call Yet Another Reaction Program (YARP), and is conceptually similar to the approach developed by Suleimanov et al.[43,44] However, these features of YARP are relatively general and could be implemented in many other emerging automated network characterization algorithms. To evaluate the performance of YARP with respect to predicting diverse chemical reactions, we have benchmarked its performance on predicting single-step reactions for 20 organic compounds compiled by Zimmerman.[32] To establish the performance of YARP on a network prediction task, we have also characterized the unimolecular decomposition of 3-hydroperoxypropanal, a small γ-ketoperoxide, for comparison with the recent benchmark of Grambow et al.[44] In both tasks, we find that YARP is capable of recapitulating prior reactions with near-perfect fidelity, discovers many new kinetically relevant pathways, and exhibits orders of magnitude reduced computational cost.

2 Methods

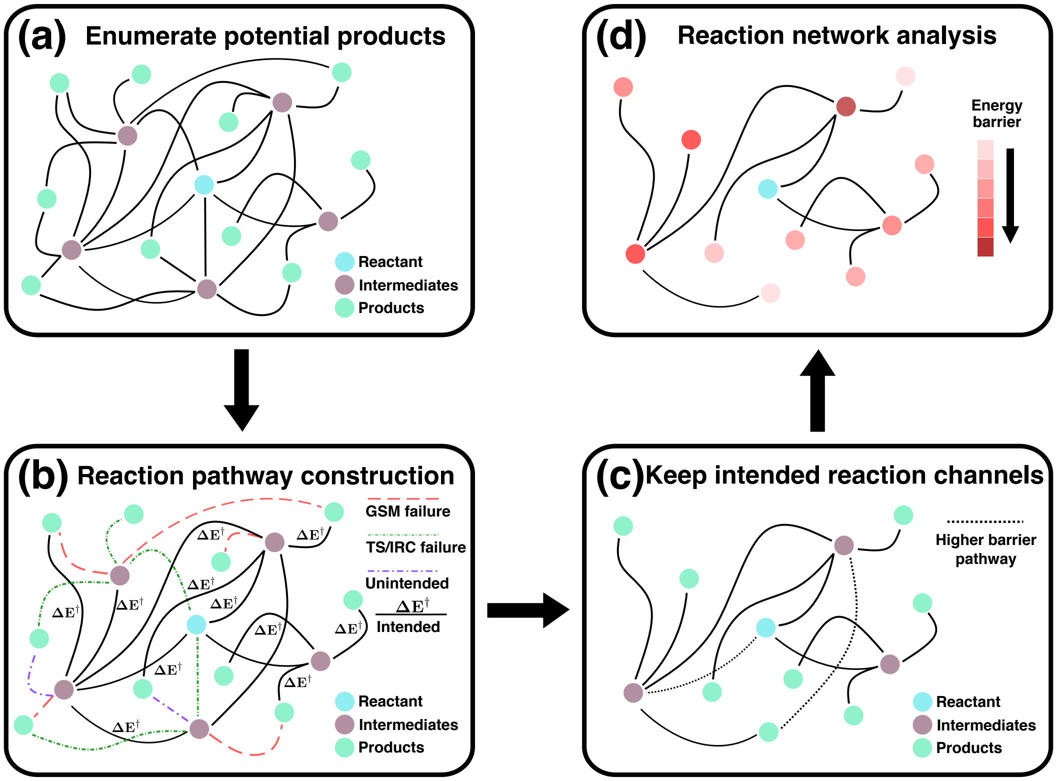

Figure 1: Overview of the YARP methodology. (a) Potential products are enumerated using ERS(s). (b) Reaction pathways are determined by applying GSM at GFN2-xTB level, Berny optimization and an IRC calculation for validation. (c) Removal of unphysical reactions based on transition state failures or unintended transition states. (d) Reaction network analysis is performed to identify the lowest activation energy pathway and determine the thermodynamic and kinetic relevance of each product.

YARP consists of three components for automatically characterizing reaction networks without the use of domain heuristics: graph-based product enumeration, reaction network construction, and reaction network analysis. A detailed description of each component is provided in the subsequent sections and an overview schematic is shown in Figure 1. In brief, potential products are recursively generated for user-supplied reactants using an automated reaction enumeration scheme and a molecular graph formalism. An initialization procedure is developed for generating three-dimensional geometries of the product and reactant states from their graph that are then used to seed a transition state search using the growing string method (GSM) coupled with the GFN2-xTB semi-empirical model chemistry. All reactions generated in this way are subsequently characterized using Berny optimization to locate transitions states at the DFT level and intrinsic reaction coordinate (IRC) calculations to validate that the identified transition states correspond with the putative reaction. At each iteration, this procedure yields a range of products, a subset or all of which can be used as inputs for additional reaction prediction. In the current work, the accuracy of single-step and multi-step recursion is evaluated on benchmarks of organic chemistry reactions and a unimolecular decomposition network, respectively.

2.1 Graph-Based Product Enumeration

Identifying transition states is the most time-consuming step in the reaction prediction process. The transition state is located in a $3N - 6$ dimensional space, where N is the number of atoms in the system (e.g., potentially including solvent atoms and catalysts). For all but the smallest systems, brute force optimization is impossible and it is critical to reduce the size of this search. Thus, several heuristics have been developed to identify transition states, either based on reaction templates, artificial forces, or using the local curvature of the potential energy surface.[36,37,46,47,50,61–65] Separately, several double-ended searching algorithms have been developed for situations where the start and end points (i.e., reactant and product states, respectively) are known in advance. In these cases, the transition state search can be recast as a one dimensional search, whereby a string (or band) of states connecting reactant(s) and product(s) is optimized to locate the transition state. In the context of reaction prediction, product enumeration is extremely inexpensive compared with transition state characterization, so one strategy that has been utilized is to enumerate putative products that are in turn efficiently characterized by one or more double-ended transition state algorithms. In this scenario, the double-ended transition state characterization algorithm carries the responsibility of discriminating between plausible and implausible reactions. In practice, however, transition state convergence failure is extremely common (e.g., $\sim 60\%$ being recently reported),[44] especially in scenarios with multiple transition states between the reactant and product structures,[34,41] suggesting that strategies to improve product enumeration are critical to ensure transition state convergence.

Elementary Reaction Step. YARP uses the concept of an elementary reaction step (ERS) to enumerate potential products and reduce the occurrence of products separated by more than one transition state.[66] Here, an ERS is defined as the smallest number of changes in bond/electron state of the reactants that results in a product with a stable Lewis structure. For neutral closed-shell organic systems, the simplest reaction step that yields non-trivial products is breaking two bonds and forming two bonds (b2f2). For instance, a "form one bond" step (f1) is inapplicable to closed-shell organics. Similarly, a b1f1 step simply reproduces the reactant(s), and a b2f1 step results in an unstable Lewis structure. Thus, b2f2 yields the most elementary set of bond rearrangements amongst closed-shell organic reactant molecules that can produce distinct products. In recent work, other groups have also included b3f3 and b4f4 steps during enumeration.[41–44,67] We note, that all b3f3 and b4f4 products can be obtained by repeated applications of b2f2 steps, although the reaction pathways may be distinct. In the case of a sequential reaction, b2f2 reaction pathways are better conditioned to double-ended searches because they reduce the possibility of transition state convergence failure due to multiple intervening transition states. For instance, out of the converged b2f2 reactions investigated here, 98.5% exhibit only one transition state due to the relative rarity of forming a stable intermediate with a single broken bond. However, in the case of a b3f3 concerted reaction (e.g., Diels-Alder), sequential application of b2f2 would discover the appropriate product, but would fail to identify the concerted mechanism. We note that the specific definition of an

ERS is system specific, and for metal-containing, ionic, or radicals systems, additional steps like b2f1, and charge transfer steps may be applicable ERS(s). More general ERS definitions are left for future work, but as we demonstrate below, even this simple ERS is surprisingly effective at comprehensively enumerating kinetically and thermodynamically relevant products of closed-shell organics.

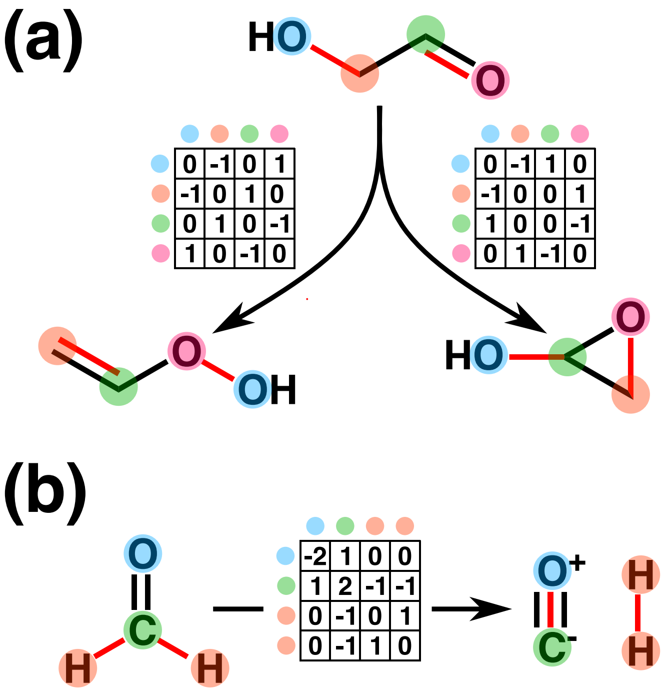

Figure 2: Illustration of two cases of "breaking two bonds and forming two bonds (b2f2)" elementary reaction step. (a) The two bonds involved in the ERS connect four different atoms. (b) An atom is shared between the two bonds involved in the ERS.

Bond-Electron Matrix Formalism. b2f2 product enumeration is implemented in YARP using the bond-electron matrix formalism developed by Ugi.[68] This approach provides a machine-readable grammar for expressing the Lewis structure of products and reactants, and encoding reactions as matrices. A bond-electron matrix, **A**, for reactant(s) with N atoms is N-dimensional and symmetric. The diagonal elements A_{ii} correspond to the number of lone valence electrons on each atom and the non-diagonal elements A_{ij} correspond to the number of covalent bonds between atoms i and j. We will refer to the bond-electron matrices corresponding to reactants and products as $\mathbf{A}$ and $\mathbf{B}$, respectively. Multiple reactant or product molecules can be combined in one bond-electron matrix, with a separate connected subgraph for each molecule. Additionally, the matrix formalism can be equivalently recast using lists to avoid handling sparse matrices, but here the matrix formalism will be retained for clarity of description.

The matrix $\mathbf{R} = \mathbf{B} - \mathbf{A}$ describes the changes in bond order and electron transfers associated with the reaction $\mathbf{A} \rightarrow \mathbf{B}$. $\mathbf{R}$ can be decomposed into a summation of two matrices, $\mathbf{R} = \mathbf{U} + \mathbf{V}$, where $\mathbf{U}$ is a diagonal matrix that describes electron transfer associated with the reaction, and $\mathbf{V}$ is a symmetric matrix that describes bond formation and dissociation.[1] In the context of the current work, the b2f2 ERS is used to generate all $\mathbf{R}$ matrices for a given set of closed shell organic reactants defined by an $\mathbf{A}$ matrix. All $\mathbf{R}$ matrices corresponding to b2f2 steps are enumerated by looping over the unique pairs of bonded atoms in $\mathbf{A}$ (i.e., the atoms involved in a double bond do not constitute a valid pair on their own). Without loss of generality, assuming the pair is composed of atoms (a,b) and (c,d), a bond is broken between each pair of atoms by setting $V_{\mathrm{ab}} = V_{\mathrm{ba}} = -1$ and $V_{\mathrm{cd}} = V_{\mathrm{dc}} = -1$ (hereafter, we do not explicitly show the symmetric terms). These broken bonds create the possibility for forming two new pairs of bonds between either (a,c) and (b,d), or (a,d) and (b,c), respectively. Thus, for each pair of two broken bonds, two distinct reaction matrices are formed with $V_{\mathrm{ac}} = V_{\mathrm{bd}} = +1$ or $V_{\mathrm{ad}} = V_{\mathrm{bc}} = +1$, respectively (Fig. 2a). In the case where an atom is shared between the two bonds [i.e., (a,b) and (a,c) compose the unique bonded pairs of atoms] we only generate a reaction matrix if atom “a” is bonded to another atom “d” that possesses at least one lone pair of electrons. In this case, the b2f2 step yields an effective bond rearrangement of $V_{\mathrm{ab}} = V_{\mathrm{ac}} = -1$ and $V_{\mathrm{bc}} = V_{\mathrm{ad}} = +1$ with a transfer of valence electrons corresponding to $U_{aa} = +2$ and $U_{dd} = -2$, and the resulting product will have formal charges on the “a” and “d” atoms. One example of such a reaction is the decomposition of formaldehyde to carbon monoxide and hydrogen (Fig. 2b). The matrix $\mathbf{B}$, corresponding to the products of each reaction, is obtained by adding the **R** matrix generated at each iteration to the reactant matrix **A**.

Finally, we note that the matrix representation is non-unique due to graph isomorphism. This leads to potential redundancy during recursive enumeration, where multiple identical reactants or products are separately subjected to reaction evaluation while in fact representing identical molecular structures. To avoid this issue, YARP implements a canonicalization procedure to sort the indexing of the bond-electron matrix based on a connectivity hash function for each atom. This hash function incorporates the elemental identity of each atom, its bonded neighbors, and their bonded neighbors, out to an arbitrary depth (e.g., 10 in the present work, but this could be set larger for more complex systems). This procedure ensures that the same reactants or products, regardless of atom indexing, resolve to the same bond-electron matrix, and thus avoids redundant calculations. Additionally, symmetry equivalent atoms evaluate to an identical hash function, which can in principle be used to further reduce the number of reactions performed. In the current study, symmetry was not used to reduce the number of reactions; thus in some cases, formally equivalent reactions (e.g., a reaction involving one of two equivalent hydrogens on a methylene) were performed multiple times.

Geometry Initialization. The bond-electron formalism provides an effective machine-readable grammar for describing reactions in terms of molecular graphs, however these graphs must be converted to three-dimensional structures for determining the transition states and thermodynamics of the reactions. Moreover, double-ended transition state algorithms are highly sensitive to the geometry of the initial structures, with documented convergence failures if the reactant and product structures are poorly aligned.[29,69] We have also observed that even in cases where transition state convergence occurs, the initial geometry can strongly impact whether the discovered transition state corresponds to the intended reaction after inspection by an IRC calculation. Within YARP, this issue is addressed by jointly optimizing reactant and product geometries for each investigated reaction. This is accomplished by first generating an optimized product geometry using the Universal Force Field (UFF)[70] as implemented in Open Babel.[71] The reactant geometry is then generated using the product geometry as a starting point, but with optimization occurring on the UFF potential energy surface corresponding to the bond-electron matrix of the reactant. The UFF optimized reactant and product geometries are then optimized at the GFN2-xTB level. Finally, root mean square deviation (RMSD) minimization, as implemented in the Atomic Simulation Environment (ASE),[72] is applied to each product geometry by rotating and translating atoms to best match the reactant.[73] This procedure yields high overall transition state convergence and a high rate of discovering transition states corresponding to intended reactions (*vide infra*). We also note that conformational sampling of the product geometries is an obvious extension of this procedure which could improve the quantitative details of discovered transition states. Given that the discovered transition states compare favorably with prior benchmarks, this has not been implemented here and will be revisited in future work.

2.2 Reaction Network Construction

YARP incrementally builds the reaction network for a given set of reactants by alternating between product enumeration and transition state characterization steps. For exploring deep networks, products from one generation may serve as reactants for a subsequent iteration of reaction enumeration. Depending on the application, it may also be necessary to include bimolecular reactions between products produced at different levels of the reaction network. In the benchmark systems (described below), we have used to YARP to perform comprehensive enumeration of b2f2 products up to two iterations. Comprehensive b2f2 sampling scales as $\binom{N}{2}$ where N is the number of unique bonded pairs of atoms in the reactant matrix. In cases with symmetry equivalent atoms the actual unique number of reactions may be substantially reduced. This scaling yields a tractable number of reactions to investigate at each iteration of product enumeration, even for relatively large reactant matrices. However, the actual performance is ultimately determined by the ability to localize accurate transition states for each reaction.

After the product enumeration phase, YARP uses one or more double-ended search algorithms to localize transition states for each reaction, followed by Berny optimization and an IRC calculation to identify and characterize the final transition state. Recently, Grambow, et. al compared several double-ended search methods, including the freezing string method (FSM), the growing string method (GSM), and the single-ended growing string method (SSM). Based on their study, GSM exhibits the best balance between computational cost and convergence rate. Although GSM is faster than SSM, directly applying it at the DFT level is still prohibitively costly for exploring deep networks or reactions involving large molecules. To decrease the computational cost in YARP, the current set of reaction pathways were localized using the GSM algorithm with the semi-empirical GFN2-xTB model chemistry prior to searching for the final DFT-level transition states via Berny optimization. This choice was motivated in part by the recent demonstration of Dohm et. al that combining GFN2-xTB with GSM to localize organometallic transition states achieved convergence for approximately 90% of the investigated reactions.[74] Here, the GSM calculations were performed by interfacing YARP with the pyGSM package[75,76] using default convergence hyperparameters (e.g., nine nodes, climbing image, and translation-rotation-internal coordinate system). After convergence of the GSM pathway, YARP automatically selects the highest energy node as a preliminary transition state candidate that is then used as a starting structure for Berny transition state optimization. After convergence, an IRC calculation is performed to determine whether the detected saddle point corresponds to the input reaction channel. If the end nodes obtained by the IRC match the input reactant and product bond-electron matrices, the reaction channel is classified as ”intended”, otherwise it is classified as ”unintended”. The latter reactions are removed from the final network, since they represent a failure of the algorithm to identify a relevant transition state either due to the fact that the attempted reaction is unphysical or because the double-ended search was poorly conditioned. In the present study, YARP used Gaussian 16 as the reference quantum chemistry engine for these characterizations.[77]

2.3 Reaction Network Analysis

The large number of reactions generated by automated algorithms creates an interpretation bottleneck for extracting mechanistic information about reactions and competing pathways. In particular, *ad hoc* evaluation of pathways and transition states is both error-prone and too costly to comprehensively evaluate large networks. To address this, we have implemented three analysis routines within YARP as a starting point to automate the extraction of semantic information from algorithmically generated reaction networks. First, rate-limiting steps are evaluated for all products identified by YARP to estimate the kinetic relevance of various reaction pathways discovered in the network. Here, the rate-limiting step is defined as the reaction with the maximum activation energy in a pathway connecting the starting reactants to a given product (Fig. 1d). For a network composed of multiple enumeration steps, multiple distinct pathways will potentially exist that yield a specific product. To compare these competing pathways, YARP performs a breadth-first search from each product, using the directed graph of reactions in the network to identify distinct pathways and keeping all pathways that terminate at the reactant node. For each product, all pathways out to $m + n$, are enumerated, where m is the depth of the product in the network (i.e., the reactant is considered as depth zero, and the product depth is defined based on the smallest number of reactions connecting it to the reactant) and n is a user-defined parameter that controls the maximum number of reactions in a pathway (e.g., taken to be two in the current work, but could be set larger or made adaptive for more complex networks). The pathway with the lowest activation energy rate-limiting step is considered the dominant pathway, and this barrier ($\Delta \mathrm{G}^{\dagger}_{\mathrm{min}}$) is reported for all products in the network. In addition, we also report the heat of reaction for each product ($\Delta\Delta H_{\mathrm{f}}$) calculated with respect to the reactants to characterize the thermodynamic relevance of the various products identified by YARP. The heat of formation of each product was calculated at the DFT level based on difference in thermal enthalpies between the reactants and the products. Finally, we have implemented an algorithm to characterize the occurrence of bond-breaking and bond-formation steps relative to the transition state. This algorithm uses a distance-based criteria to define changes in bond configuration. Specifically, bond breaking (formation) is defined to occur when the observed separation between a pair of atoms is greater (less) than 1.2 times the summation of the UFF radii of the atoms. This algorithm is applied to the minimum energy pathway generated by the IRC calculation to determine which events are associated with the transition state, and whether they occur sequentially or in a concerted fashion. The specific position within the minimum energy path of the identified bond changes is relatively insensitive to the choice of scaling factor, but it cannot be set too large without detecting spurious bonds. In the case of double-bonds, the distance based criteria is inapplicable and the algorithm assigns bond-breaking to occur at the same time as a bond-forming event involving either of the double-bonded atoms.

2.4 Performance Statistics

The number of quantum chemistry gradient calls is a commonly used surrogate for the computational cost of a reaction exploration method, as the transition state calculations represent the most expensive step in reaction characterization.[32,43,44,78] Following this approach, we have reported the distribution of DFT gradient calls associated with all reaction channels explored with YARP. Since the cost of each gradient call at the GFN2-xTB level is negligible compared with the cost at the DFT level, gradient calls associated with GFN2-xTB pre-optimization are not included in these totals. Likewise, the gradient calls associated with IRC calculations are also excluded from these totals to be consistent with the gradients statistics reported by Grambow et al., as the IRC calculations are a validation activity and not directly associated with finding the transition states.

Additionally, statistics associated with the success rate of transition state convergence and whether transition states correspond to intended channels are reported with the following definitions. The success rate corresponds to the fraction of unique reactions that completed all calculations (i.e., GSM, Berny, and IRC) compared with the total number of unique reactions attempted. The intended rate corresponds to the fraction of unique reactions that completed all calculations and exhibit an intended transition state compared with the total number of unique reactions attempted. Thus, the intended rate is bound from above by the success rate. Note, there is a minor ambiguity in how these rates have been defined in previous work, depending on whether one uses the total number of reactions or the total number of unique reactions in calculating these rates. The distinction is that the same product may be obtained by multiple single-step reactions (e.g., if there are symmetry equivalent atoms in the reactant or in the product) and we are here defining "unique" reactions to be reactions that yield distinct products. We adopt the convention that if at least one reaction out of a set of equivalent reactions is classified as successful (intended), then that unique reaction is classified as successful (intended). Our rationale is that since at least one transition state was successful (intended) for such a reaction, any discrepancy in the rates calculated using unique reactions versus total reactions reflects the sensitivity of converging transition states rather than the feasibility of the reaction. We report rates on a unique reaction basis within the main text for each distinct reactant investigated in this study. We also report statistics on a per reaction basis in the SI (Fig. S2).

2.5 Benchmark systems

In the present work, YARP was used to investigate two classes of reaction prediction problems. The first was predicting a set of single-step organic reactions curated by Zimmerman.[32] The Zimmerman dataset was developed to validate the performance of GSM and not as a test set for reaction exploration. Here, a comparison between the lowest barrier pathways predicted by YARP and the Zimmerman reactions was used to validate YARP's ability to (re)discover established reactions across a range of organic chemistries without domain knowledge. The Zimmerman dataset includes 25 compounds and 105 single-step reactions. We excluded ionic species (NH_3BH_3 and the taxadiene carbocation) and reactants with incomplete octets (NH_2BH_2 and SiH_2) as they require a more general set of ERSs beyond b2f2. We also excluded alanine dipeptide, since Zimmerman only reported a conformational rearrangement, rather than a reaction for this compound. The final dataset evaluated here was composed of 20 distinct reactants and 61 distinct reactions. YARP was used to characterize all b2f2 reactions for each reactant, resulting in a total of 533 distinct reactions. The only molecule where all b2f2 reactions were not performed was the phenyl ether, for which the default setting of YARP is to exclude breaking bonds in benzene rings. The Berny optimization and IRC calculations were performed at the B3LYP/6-31G** level to be consistent with the prior work.

YARP was also used to characterize the thermal decomposition network of 3-hydroperoxypropanal, recently benchmarked by Grambow et al.[44] In this earlier work, the reaction networks predicted by five distinct reaction discovery algorithms were compared, in total consisting of 55 distinct products and 75 intended reactions. YARP was used to perform two iterations of b2f2 reactions for this system (i.e., all b2f2 reactions for 3-hydroperoxypropanal and the first generation of products) to evaluate the performance of YARP with respect to computational cost and reaction coverage. b2f2 reactions were only applied to products of 3-hydroperoxypropanal whose transition state calculations converged to the intended reaction as determined from IRC calculations. For the general problem of generating a deep reaction network, recursive application of b2f2 steps to all products at each iteration is typically unnecessary as kinetic modeling can be used to prioritize which products are capable of further reaction.[23,79,80] However, the decision to investigate all b2f2 reactions out to the second generation for the 3-hydroperoxypropanal network was made to provide a direct comparison with the earlier study, which reported reaction channels up to b4f4. The Berny optimization and IRC calculations were performed at the B3LYP/6-31+G* level to be consistent with the prior work.

3 Result and Discussion

3.1 Single-Step Reaction Prediction Benchmark

The trade-off between reaction exploration and computational cost represents a major limitation on the breadth and depth of reaction networks that can be discovered by automated algorithms. Within the context of YARP, we have adopted a generic b2f2 ERS that yields a tractable number of reactions to characterize and which we reasoned would be able to recapitulate typical closed shell organic reactions without being biased solely towards known reactivities. To establish the effectiveness of the b2f2 ERS, we applied YARP to 20 distinct reactants curated by Zimmerman[29] and compared the established reactions with the low-barrier channels predicted by YARP (Fig. 3).

First, we note the low number of DFT-level gradient calls required by YARP to identify the transition states of the attempted reactions (Fig. 3a). We observe the range of gradient calls is from 4 to 50 with an average of ~ 13 per reaction and around 75% of reactions needed less than 20 gradient calls. In comparison, Zimmerman reports on average 468 gradient calls for his most robust set of GSM hyperparameters (i.e. 11 images with climbing image optimization and overlap criterion) and approximately 380 gradient calls on average when using nine images as employed here.[32] Although the reference datasets have minor differences as described in the Benchmark Systems section, YARP clearly exhibits a qualitative reduction in computational cost compared with direct GSM localization at the DFT level.

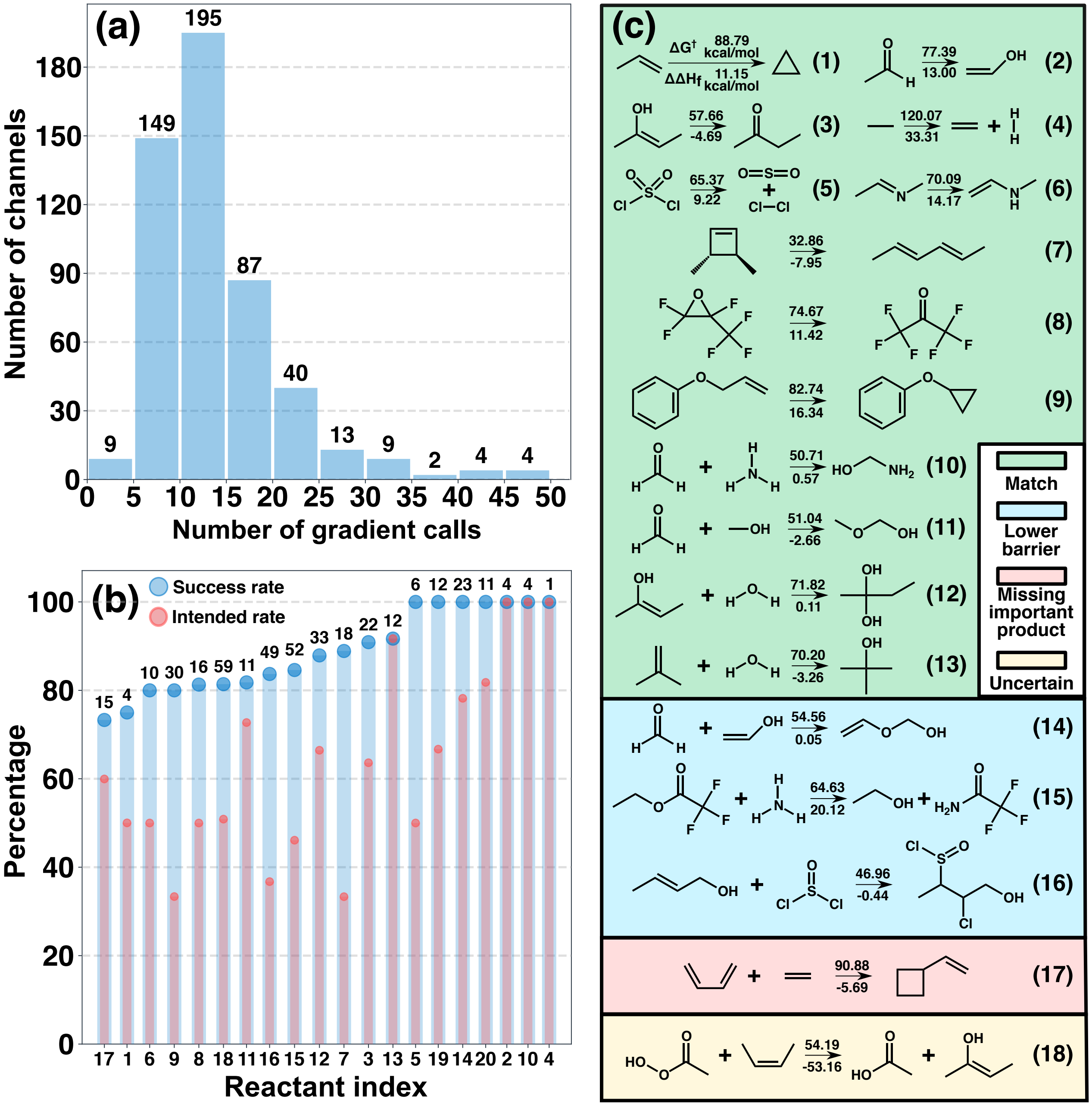

Figure 3: Overview of YARP performance on predicting reactions from the Zimmerman dataset. (a) The distribution of gradient calls required to converge each intended reaction, with labels over each bar indicating the number of reactions in each bin. (b) The success and intended rates of unique reactions involving each reactant. (c) The minimum activation energy pathways discovered by YARP, classified based on comparison with Zimmerman's predictions.

A dramatic reduction in gradient calls is moot if YARP cannot also successfully localize transition states. Thus, in Figure 3b we have evaluated the transition state success and intended rates averaged across all unique reactions involving each reactant. The success rate varies between 75% and 100% with an overall average of 86.2% for all 20 reactants. The intended rate varies between 33.3% and 100% with an average of 63.3%. The comparison between the success rate and the intended rate illustrates that the latter depends on reactant complexity, while the success rate is consistently high. In particular, the number of potential products and the dimensionality of the potential energy surface increase with reactant size and complexity, which also increases the occurrence of unphysical reactions. For some simple systems, like R2, R4 and R10, where the number of heavy atoms are 3, 2 and 3, respectively, both the success rate and intended rate reach 100%. However, for R7, a stable four-membered ring, and R9, an aromatic compound with 10 heavy atoms, the intended rate is only 33.3% while the success rate is still over 80%. We note that even with a comprehensive transition state search, neither the success rate nor the intended rate would necessarily reach 100% when comprehensive reaction enumeration is being used. Because the transition state algorithm within YARP is playing the role of discriminating between physical and unphysical reactions, it may be that some of the enumerated reactions are poorly conditioned and no physical transition state exists that directly connects those reactants and products. For instance, considering only the reactions that were characterized by both Zimmerman and YARP, YARP exhibits a success rate of 100%, which is the same as in the earlier work. The intended rate was not reported by Zimmerman, but it is similarly high in YARP at 98% for this subset of reactions.

The goal of utilizing this large test set, which includes a broad range of standard organic reactions and functional groups, was to validate that YARP is capable of (re)discovering diverse reactions using a generic ERS. To show the kinetically favorable reactions discovered by YARP, the reaction with the lowest activation energy for each reactant was selected and compared with the Zimmerman dataset (Fig. 3c). The green region includes 13 reactions which are the lowest energy barrier reactions in both the Zimmerman dataset and YARP. For the three reactions in the blue region, YARP identified the same reactions as Zimmerman, but also discovered reactions with even lower activation energy that were not identified in the previous study. For R17 and R18, in the red and yellow regions, respectively, the lowest activation energy reactions in the Zimmerman dataset are b3f3 reactions that are not discovered by single-step b2f2 enumeration in YARP. In particular, R17 represents the Diels-Alder reaction, which is a concerted b3f3 reaction that would not be discovered even by repeated application of the b2f2 ERS. In contrast, the lowest energy barrier reaction of R18 predicted by YARP was not investigated by Zimmerman, which makes it uncertain which reaction is in fact more likely. Additionally, Zimmerman reports only one reaction pathway for R19 (Dimethylphosphine + ethene) and R20 (Trimethylphosphine + Oxirane) and neither is a b2f2 reaction step and thus cannot be compared here. To summarize these comparisons, YARP was able to automatically discover all of the previously reported b2f2 reactions and identify them as important channels among all investigated reactions. In addition, YARP discovered several lower activation energy reactions that were not previously reported. On the other hand, all reactions that were missed by YARP can be rationalized by the fact that they either required repeated applications of the b2f2 ERS, which was not pursued here, or the inclusion of complementary ERS(s). In the latter case, we note that such reactions are relatively contextual (e.g., atoms capable of expanded octets or multiple double bonds), and it may be desirable to leverage additional elementary chemical information available from the reactant graph when expanding the scope of ERS(s).

Reaction Mechanisms Discovered by YARP. The rationale behind using ERS(s) is to better condition the transition state localization by reducing the occurrence of reactions that possess multiple transition states. We note that across all of the reactions that were successfully localized in this study, 98.5% exhibit only a single transition state in the minimum energy pathway

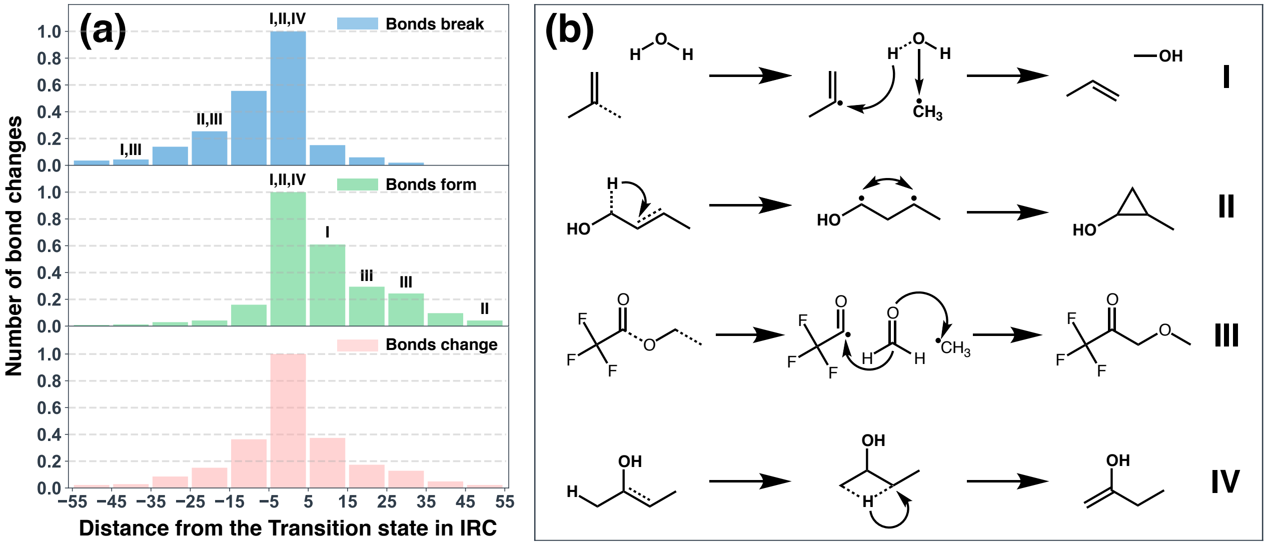

Figure 4: Characterization of sequential and concerted reaction mechanisms discovered by YARP. (a) The histogram of when bond breaking (top), bond forming (middle), or any bond change (bottom), occur relative to the transition state based on analysis of the IRC minimum pathway. (b) Representative sequential (I,II,III) and concerted reactions (IV), corresponding to the labels in (a).

determined by GSM. Nevertheless, the reactions that are discovered by YARP still exhibit a broad range of mechanistic diversity with respect to when bonds break and form relative to the transition state (Fig. 4). Based on the bond-breaking and bond-forming histograms in Figure 4a, we observe the intuitive result that bond-breaking events skew earlier than the transition state and bond-forming events skew later than the transition state. Similarly, the aggregated bond-breaking and bond-forming distributions are centered about the transition state, indicating that the transition state of most reactions consists of at least one change in bond order (i.e., more than 80% of bond changes occur within ±20 steps of the transition state). Additionally, we can further distinguish between several types of sequential reactions and concerted reactions that are predicted by YARP (**I-IV** in Fig. 4b). Among the sequential reactions, we observe at least one bond-breaking event before the transition state in all cases, followed by b1f2 type transition states (**I**), b1f1 transition states (**II**), and transition states only associated with conformational rearrangement (**III**). Typical examples of **I** are bimolecular reactions, where the reaction can proceed in mainly concerted fashion after an initial bond-breaking event. Typical examples of **II** are reactions involving ring closures, where the transition state is associated with a partial bond rearrangement, followed by a later ring closure. Typical examples of **III** are reactions that pass through a multi-molecular structure, where a significant conformational change must occur before executing the bond formation steps. In contrast, we also observe a high occurrence of concerted reactions (**IV**), defined here by reactions where all bond rearrangements occur within ±20 steps of the transition state. For example, in the hydride shift reaction shown in Figure 4b, all bond changes occur within five steps of the transition state. Although the classifications of specific reactions are subject to how many images are included in the transition state region, these results demonstrate that in addition to being able to discover chemically diverse reactivities using ERS(s), YARP also discovers a broad range of reaction mechanisms in the predicted pathways. The automatic classification of reaction mechanisms as performed here could also be extended to include more fine-grained mechanistic classification that could in turn inform ERS definitions.

3.2 Unimolecular Degradation Network of 3-hydroperoxypropanal

In addition to single-step reaction discovery, automated reaction prediction has potentially the largest relevance to elucidating complex multi-step reaction networks. For resolving potentially deep and/or broad networks, reducing the computational cost while still robustly identifying all kinetically relevant pathways remains an outstanding problem within the field. Additionally, a dearth of benchmark systems for which extensive reaction networks, including transition states, predicted products, and gradient call statistics have been reported poses a major challenge to unequivocally establishing the performance of competing algorithms. Recently Grambow et al. published a useful benchmark study on the unimolecular decomposition of 3-hydroperoxypropanal that provides a comparison of network predictions for five distinct reaction discovery algorithms.[44] This work

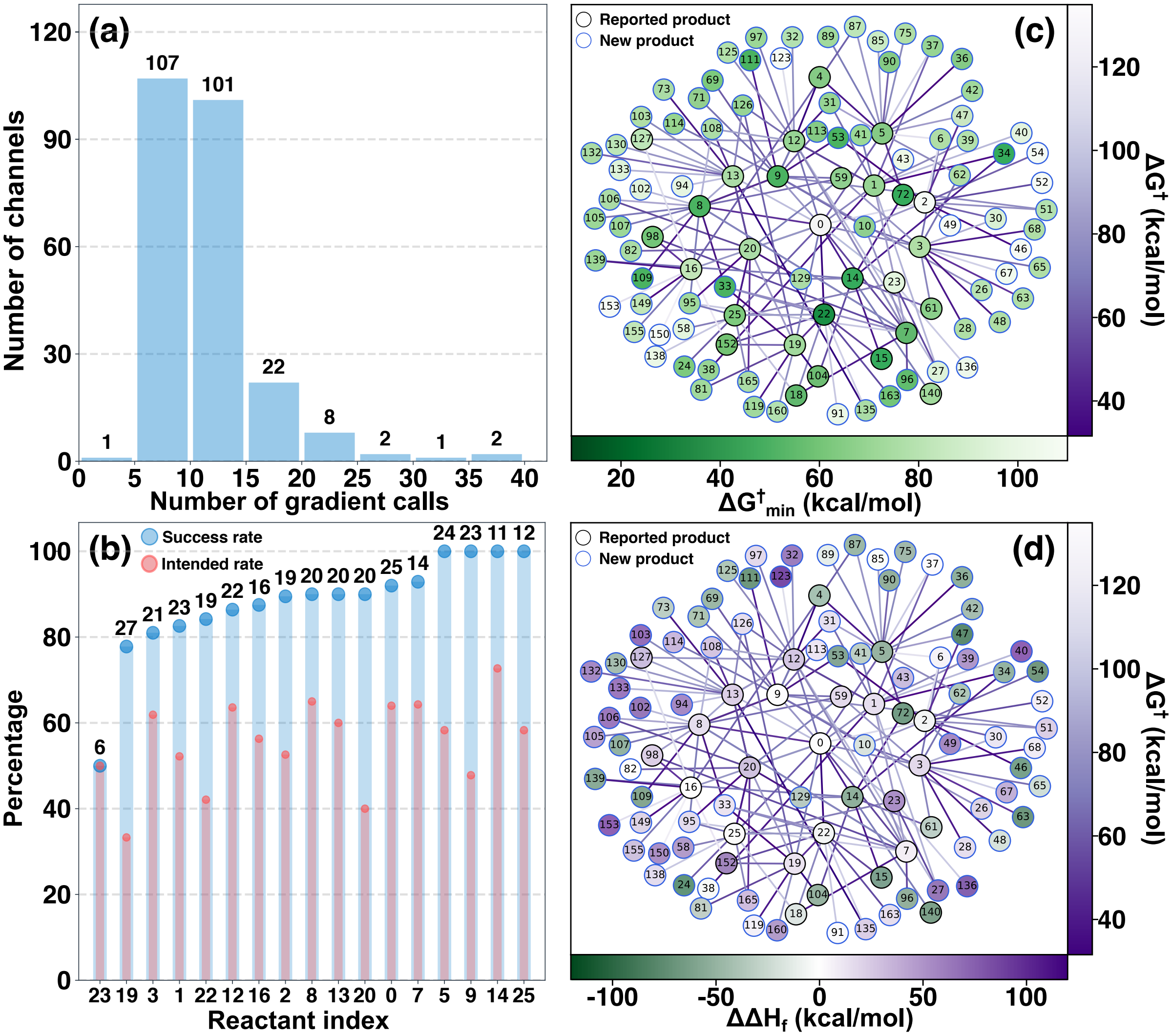

Figure 5: Overview of YARP performance on predicting unimolecular degradation of 3-hydroperoxypropanal. (a) The distribution of gradient calls required to converge each intended reaction, with labels over each bar indicating the number of reactions in each bin. (b) The success and intended rates of unique reactions involving each reactant. (c) Degradation network with activation barriers for rate-limiting steps, which represents kinetic accessibility. (d) Degradation network with heat of reaction, which represents thermodynamic accessibility.

was also followed by a report by Maeda et al. providing updated results for the single-component artificial force induced reaction method (SC-AFIR).[78] 3-hydroperoxypropanal is a representative γ-ketohydroperoxide, an important class of molecules for autooxidation and pre-ignition chemistries with complex reactions sequences and intermediates that are still under active investigation. In the context of the current work, 3-hydroperoxypropanal served as a useful benchmark system for evaluating the performance of YARP with respect to multi-step reaction networks and computational cost.

To provide an illustration of the reaction filtering that is common in network exploration, we first used YARP to recursively enumerate all b2f2 products out to two reaction steps. Before transition state characterization, this yields a putative reaction network consisting of 286 vertices (reactants and products) and 1148 edges (distinct reactions). We then excluded products involving three-membered and four-membered rings due to the fact that these have a relatively high heat of formation and are unlikely to be kinetically relevant. After removing these compounds, the reaction network consisted of 174 vertices and 539 edges. After performing transition state characterizations, IRC calculations, and retaining only products connected by intended channels, the final network is composed of 107 vertices and 173 edges (Fig. 5).

For comparison, in the earlier work of Grambow et al., reaction products out to b4f4 type rearrangements were explored, consisting of 562 reaction channels. Of these, 75 intended channels were discovered involving 55 distinct products (aggregated across all five of the methods), which can be compared with the YARP results above. As applied in this case, YARP attempted to identify reaction pathways yielding all degradation products of 3-hydroperoxypropanal up to b4f4 type rearrangements, but excluding compounds containing three-membered rings, four-membered rings, or ionic species due to the selected ERS and filtering. That is, YARP was used to characterize a similar number of reactions and has not been run to a deeper level of recursion to artificially inflate the number of discovered reactions and products. Thus, differences between the YARP network and that of Grambow, beyond those noted in the previous sentences, are due to the improved localization of intended reaction channels by YARP.

Figure 5c shows the reaction network predicted by YARP, including previously identified products and newly discovered products. Less products containing small rings and ionic species, YARP automatically discovered all previously reported products and managed to connect them by intended channels to the 3-hydroperoxypropanal reactant. In addition, YARP discovered 77 new products and 157 new reactions that have not been previously reported. Constructing this network with YARP required only 8364 DFT level gradient calls. The distribution of number of gradient calls per intended channel is shown in Figure 5(a), which illustrates that more than 85% of the intended channels required less than 15 DFT gradient calls. In comparison, directly applying GSM at DFT level as reported in the earlier benchmark required 756227 gradient calls to explore 562 reaction channels, representing a nearly 100-fold reduction for YARP. In addition, the transition state success and intended rates averaged across all unique reactions involving each reactant (i.e., 3-hydroperoxypropanal and all intended b2f2 products of 3-hydroperoxypropanal comprise the reactants for this network) are shown in Figure 5b. We observe that the success rates for all reactants besides species 19 and 23 are larger than 80%, while the intended rate varies between 33.3% and 72.7%. The average overall success and intended rates calculated on a unique reaction basis (see methods for the distinction) for this network are 88.8% and 54.7%, respectively. For comparison with the earlier benchmark, we have also calculated the success and intended rates on a per attempted reaction basis (Fig. S2), which yields overall success and intended rates of 81.4% and 41.6%, respectively, for YARP. In contrast, Grambow et al. reported success and intended rates on a per attempted reaction basis for their GMS based reaction discovery algorithm of 38% and 4%, respectively.[44] This dramatic difference in both the success and intended rates suggests that ERS enumeration provides more realistic reaction channel candidates that are better conditioned to GSM characterization.

To evaluate whether the new products and reactions discovered by YARP are of kinetic or thermodynamic significance, we also performed several additional characterizations of the network and reaction pathways. Illustrations of the reaction network are provided in Figures 5c-d that show the activation energy associated with all reactions (edge color) as well as the activation energy of the rate-limiting step ($\Delta G^{\dagger}_{\text{min}}$) and heat of reaction ($\Delta\Delta H_{\text{f}}$) for each product. Although these illustrations compress a lot of information, it is apparent that several of the newly discovered products are both thermodynamically and kinetically favorable compared to previously reported products. For example, products 34, 53, 109, and 111 (among others) all exhibit relatively low-barrier rate-limiting steps ($\Delta G^{\dagger}_{\text{min}} < 55$kcal/mol) and strongly exothermic relationships to 3-hydroperoxypropanal ($\Delta\Delta H_{\text{f}} < -50$kcal/mol). In addition, we observe several products (6, 10, and 24) that YARP identified intended channels for, while the previous benchmark identified these products but failed to converge intended transition states. The difference in this case is that the previous benchmark only attempted a single-step reaction that failed to converge to an intended transition state, whereas YARP identified alternative multi-step pathways for these products. These results clearly demonstrate that the relatively low success and intended rates reported previously ultimately lead to incomplete reaction discovery and the neglect of potentially important reaction products. Documentation of all products, rate-limiting reaction barriers, and reaction pathways predicted by YARP is provided in the Table S3.

In addition to discovering new products, we also observe that YARP predicts new lower barrier reaction pathways for ten previously reported products (i.e., alternative reaction steps with >5 kcal/mol reduction in activation barrier compared to previous reports). For example, in Figure 6 we show five reaction pathways predicted by YARP that exhibit >20 kcal/mol reduction in activation energy compared with the previous benchmark. In all of these cases, we observe that YARP identifies a favorable multi-step reaction pathway, whereas Grambow et al. report a single-step reaction. This difference can be rationalized by the use of a b2f2 ERS by YARP in the

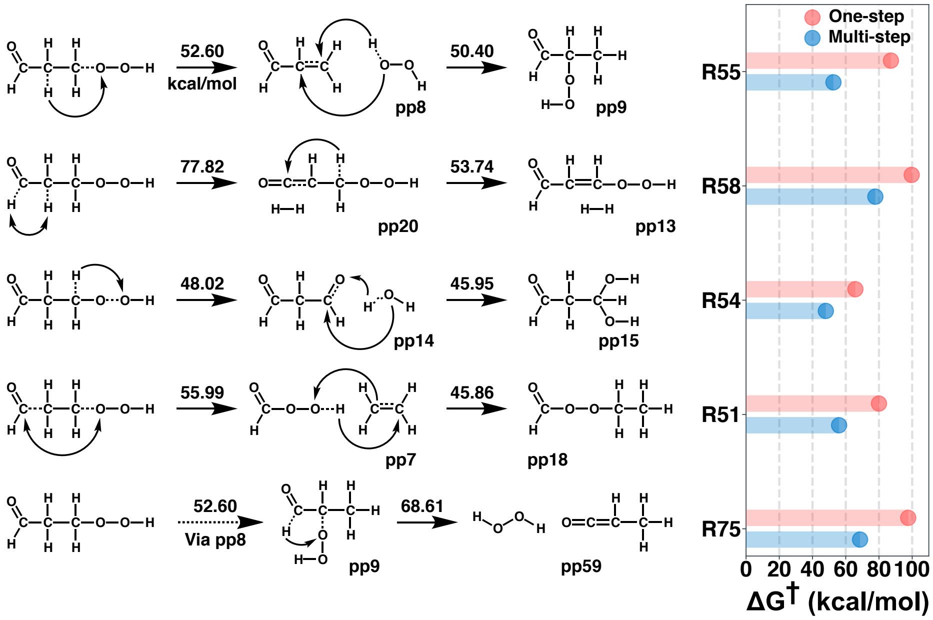

Figure 6: Five reaction pathways identified by YARP that exhibit more than 20 kcal/mol reduction in activation energy. ppN refers to Nth product identified by YARP and RX refers to corresponding index used by Grambow et al. The number above each arrow lists the activation barrier for each reaction.

present study, whereas in the earlier work, reactions up to b4f4 were attempted directly. Thus, even in cases where a more complex ERS can be successfully converged to an intended channel, it may not be the most kinetically relevant pathway. In contrast, we only observe two cases (R28 and R69, in the original benchmark) where Grambow et al. reported a lower barrier pathway for products identified by YARP (0 → 14 → 61 and 0 → 14 → 72, respectively). Both of these cases involve a b3f3 reaction that was not investigated by YARP. We stress that both studies used the same level of DFT to characterize transition states, and thus these differences reflect the ERS and geometric initialization procedure specific to YARP. In general, these results suggest that simpler ERS(s) lead to better convergence and more physically relevant pathways compared with more complex ERS(s) like b3f3 and b4f4. The general conditions under which this is true is an interesting question, and we also remind the reader of the failure of YARP to identify the Diels-Alder reaction in the Zimmerman dataset. Thus, more complex ERS(s) are definitely needed in some cases, but empirically these situations seem to be rare for closed-shell neutral organics.

4 Conclusion

The high computational cost and inconsistent reaction coverage of automated reaction prediction methodologies are still obstacles to deep reaction networks exploration. While strategies have been developed to address these problems individually, solving them simultaneously is still an ongoing challenge. In practice, a trade-off occurs in which increasing reaction coverage typically involves increasing the number of reactions evaluated at a high-level of theory, whereas decreasing cost typically involves characterizing less reactions. This has led to a bifurcation between relatively low-cost prediction methodologies that use domain heuristics to limit the scope of feasible reactions, and relatively high-cost but systematic methodologies that maximize reaction coverage. As highlighted here, even in the latter case of systematic reaction exploration, prevailing methods still miss many reactions due to inconsistent transition state convergence. However, this presentation of the trade-off implicitly assumes that capping the number of reactions is the only way of reducing cost and neglects the fact that substantial computational gains are still possible by relatively straightforward modifications to reaction enumeration, geometry initialization, and transition state convergence algorithms.

Comparing our methodology, YARP, with contemporary benchmarks, we observe that computational cost can be reduced up to 100-fold, while simultaneously improving the diversity of identified products and reaction pathways. Moreover, YARP exhibits an overall improvement of success and intended rates for identifying transition states, exhibiting up to 2-fold and 10-fold improvements, respectively, in comparison with prevailing methods. These results demonstrate that the computational limitations associated with systematic reaction prediction are largely surmountable by algorithmic improvements rather than narrowing the scope of assayed reactions. We also note that when characterizing deep networks, it is typically unnecessary to explore all branches to an equal depth, which is the common assumption in estimates of quadratic scaling of the number of reactions in a network with respect to number of intermediates. As has been pointed out by Lu, Van de Vijver and others, real reaction networks typically exhibit linear scaling in the number of kinetically relevant reactions with respect to number of intermediates. Thus, combining systematic reaction prediction, on a per reactant basis, with kinetic modeling to prioritize which branches to explore to increased reaction depth is a promising pathway towards achieving linear scaling of deep network characterizations, without sacrificing reaction discovery.

Ongoing efforts are occurring in this regard to combine YARP with kinetic modeling to elucidate deep reaction networks relevant to electrolyte chemistry and materials degradation. There are also several obvious extensions to the ERS proposed here that will extend the scope of systems that YARP can be applied to. As the data generated by high-throughput physics-based approaches like YARP mature, many opportunities will also be created to utilize machine learning approaches to further accelerate transition state characterization and refine reaction enumeration. In combination, these avenues of research make it likely that automated reaction prediction will develop into a routine research tool for elucidating currently intractable reaction network problems.

References

(1) Di Maio, F. P.; Lignola, P. G. KING, a kinetic network generator. *Chem. Eng. Sci.* **1992**, *47*, 2713–2718.

(2) Ranzi, E.; Gaffuri, P.; Faravelli, T.; Dagaut, P. A wide-range modeling study of n-heptane oxidation. *Combust. Flame* **1995**, *103*, 91–106.

(3) Curran, H. J.; Gaffuri, P.; Pitz, W. J.; Westbrook, C. K. A comprehensive modeling study of

n-heptane oxidation. *Combust. Flame* **1998**, *114*, 149–177.

(4) Westbrook, C. K.; Mizobuchi, Y.; Poinsot, T. J.; Smith, P. J.; Warnatz, J. Computational combustion. *Proc. Combust. Inst.* **2005**, *30*, 125–157.

(5) Sarathy, S. M.; Westbrook, C. K.; Mehl, M.; Pitz, W. J.; Togbe, C.; Dagaut, P.; Wang, H.; Oehlschlaeger, M. A.; Niemann, U.; Seshadri, K.; Veloo, P. S. Comprehensive chemical kinetic modeling of the oxidation of 2-methylalkanes from C7 to C20. *Combust. Flame* **2011**, *158*, 2338–2357.

(6) Hatzimanikatis, V.; Li, C.; Ionita, J. A.; Broadbelt, L. J. Metabolic networks: enzyme function and metabolite structure. *Curr. Opin. Struct. Biol.* **2004**, *14*, 300–306.

(7) Mayeno, A. N.; Yang, R. S. H.; Reisfeld, B. Biochemical reaction network modeling: predicting metabolism of organic chemical mixtures. *Environ. Sci. Technol.* **2005**, *39*, 5363–5371.

(8) Rodrigo, G.; Carrera, J.; Prather, K. J.; Jaramillo, A. DESHARKY: automatic design of metabolic pathways for optimal cell growth. *Bioinformatics* **2008**, *24*, 2554–2556.

(9) Wu, D.; Wang, Q.; Assary, R. S.; Broadbelt, L. J.; Krilov, G. A computational approach to design and evaluate enzymatic reaction pathways: application to 1-butanol production from pyruvate. *J. Chem. Inf. Model.* **2011**, *51*, 1634–1647.

(10) Stine, A.; Zhang, M.; Ro, S.; Clendennen, S.; Shelton, M. C.; Tyo, K. E.; Broadbelt, L. J. Exploring De Novo metabolic pathways from pyruvate to propionic acid. *Biotechnol. Prog.* **2016**, *32*, 303–311.

(11) Smith, G. P.; Frenklach, M.; Feeley, R.; Packard, A.; Seiler, P. A system analysis approach for atmospheric observations and models: Mesospheric HOx dilemma. *J. Geophys. Res.: Atmos.* **2006**, *111*.

(12) Centler, F.; Dittrich, P. Chemical organizations in atmospheric photochemistries—A new method to analyze chemical reaction networks. *Planet. Space Sci.* **2007**, *55*, 413–428.

(13) Jalan, A.; Allen, J. W.; Green, W. H. Chemically activated formation of organic acids in reactions of the Criegee intermediate with aldehydes and ketones. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16841–16852.

(14) Vereecken, L.; Aumont, B.; Barnes, I.; Bozzelli, J.; Goldman, M. J.; Green, W. H.; Madronich, S.; McGillen, M. R.; Mellouki, A.; Orlando, J. J.; Picquet-Varrault, B. Perspective on mechanism development and structure-activity relationships for gas-phase atmospheric chemistry. *Int. J. Chem. Kinet.* **2018**, *50*, 435–469.

(15) Rousso, A. C.; Hansen, N.; Jasper, A. W.; Ju, Y. Identification of the Criegee intermediate reaction network in ethylene ozonolysis: impact on energy conversion strategies and atmospheric chemistry. *Phys. Chem. Chem. Phys.* **2019**, *21*, 7341–7357.

(16) Corey, E. J.; Wipke, W. T. Computer-assisted design of complex organic syntheses. *Science* **1969**, *166*, 178–192.

(17) Grzybowski, B. A.; Bishop, K. J.; Kowalczyk, B.; Wilmer, C. E. The'wired'universe of organic chemistry. *Nat. Chem.* **2009**, *1*, 31–36.

(18) Skoraczyński, G.; Dittwald, P.; Miasojedow, B.; Szymkuć, S.; Gajewska, E. P.; Grzybowski, B. A.; Gambin, A. Predicting the outcomes of organic reactions via machine learning: are current descriptors sufficient? *Sci. Rep.* **2017**, *7*, 1–9.

(19) Coley, C. W.; Green, W. H.; Jensen, K. F. Machine learning in computer-aided synthesis planning. *Acc. Chem. Res.* **2018**, *51*, 1281–1289.

(20) Schwaller, P.; Petraglia, R.; Zullo, V.; Nair, V.; Haeuselmann, R.; Pisoni, R.; Bekas, C.; Iuliano, A.; Laino, T. Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy. *Chem. Sci.* **2020**, *11*, 3316–3325.

(21) Simm, G. N.; Vaucher, A. C.; Reiher, M. Exploration of reaction pathways and chemical transformation networks. *J. Phys. Chem. A* **2018**, *123*, 385–399.

(22) Green, W. H. *Computer Aided Chemical Engineering*; Elsevier, 2019; Vol. 45; pp 259–294.

(23) Vernuccio, S.; Broadbelt, L. J. Discerning complex reaction networks using automated generators. *AIChE J.* **2019**, *65*, e16663.

(24) Coley, C.; Jin, W.; Rogers, L.; Jamison, T.; Jaakkola, T.; Green, W.; Barzilay, R.; Jensen, K. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chem. Sci.* **2019**, *10*, 370–377.

(25) Schreck, J. S.; Coley, C. W.; Bishop, K. J. Learning retrosynthetic planning through simulated experience. *ACS Cent. Sci.* **2019**, *5*, 970–981.

(26) Halgren, T. A.; Lipscomb, W. N. The synchronous-transit method for determining reaction pathways and locating molecular transition states. *Chem. Phys. Lett.* **1977**, *49*, 225–232.

(27) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901–9904.

(28) Peters, B.; Heyden, A.; Bell, A. T.; Chakraborty, A. A growing string method for determining transition states: Comparison to the nudged elastic band and string methods. *J. Chem. Phys.* **2004**, *120*, 7877–7886.

(29) Zimmerman, P. M. Growing string method with interpolation and optimization in internal coordinates: Method and examples. *J. Chem. Phys.* **2013**, *138*, 184102.

(30) Birkholz, A. B.; Schlegel, H. B. Path optimization by a variational reaction coordinate method. I. Development of formalism and algorithms. *J. Chem. Phys.* **2015**, *143*, 244101.

(31) Behn, A.; Zimmerman, P. M.; Bell, A. T.; Head-Gordon, M. Efficient exploration of reaction paths via a freezing string method. *J. Chem. Phys.* **2011**, *135*, 224108.

(32) Zimmerman, P. M. Reliable transition state searches integrated with the growing string method. *J. Chem. Theory Comput.* **2013**, *9*, 3043–3050.

(33) Martínez, T. J. Ab initio reactive computer aided molecular design. *Acc. Chem. Res.* **2017**, *50*, 652–656.

(34) Dewyer, A. L.; Argüelles, A. J.; Zimmerman, P. M. Methods for exploring reaction space in molecular systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1354.

(35) Unsleber, J. P.; Reiher, M. The exploration of chemical reaction networks. *Annu. Rev. Phys. Chem.* **2020**, *71*, 121–142.

(36) Luo, Y.; Maeda, S.; Ohno, K. Automated exploration of stable isomers of H+ (H2O) n (n= 5–7) via ab initio calculations: An application of the anharmonic downward distortion following algorithm. *J. Comput. Chem.* **2009**, *30*, 952–961.

(37) Maeda, S.; Taketsugu, T.; Morokuma, K. Exploring transition state structures for intramolecular pathways by the artificial force induced reaction method. *J. Comput. Chem.* **2014**, *35*, 166–173.

(38) Maeda, S.; Harabuchi, Y.; Takagi, M.; Taketsugu, T.; Morokuma, K. Artificial force induced reaction (AFIR) method for exploring quantum chemical potential energy surfaces. *Chem. Rec.* **2016**, *16*, 2232–2248.

(39) Shang, C.; Liu, Z. P. Stochastic surface walking method for structure prediction and pathway searching. *J. Chem. Theory Comput.* **2013**, *9*, 1838–1845.

(40) Zhang, X. J.; Liu, Z. P. Reaction sampling and reactivity prediction using the stochastic surface walking method. *Phys. Chem. Chem. Phys.* **2015**, *17*, 2757–2769.

(41) Zimmerman, P. M. Automated discovery of chemically reasonable elementary reaction steps. *J. Comput. Chem.* **2013**, *34*, 1385–1392.

(42) Zimmerman, P. M. Navigating molecular space for reaction mechanisms: an efficient, automated procedure. *Mol. Simul.* **2015**, *41*, 43–54.

(43) Suleimanov, Y. V.; Green, W. H. Automated discovery of elementary chemical reaction steps using freezing string and Berny optimization methods. *J. Chem. Theory Comput.* **2015**, *11*, 4248–4259.

(44) Grambow, C. A.; Jamal, A.; Li, Y. P.; Green, W. H.; Zador, J.; Suleimanov, Y. V. Unimolecular reaction pathways of a γ-ketohydroperoxide from combined application of automated reaction discovery methods. *J. Am. Chem. Soc.* **2018**, *140*, 1035–1048.

(45) Broadbelt, L. J.; Stark, S. M.; Klein, M. T. Computer generated pyrolysis modeling: on-the-fly generation of species, reactions, and rates. *Ind. Eng. Chem. Res.* **1994**, *33*, 790–799.

(46) Gao, C. W.; Allen, J. W.; Green, W. H.; West, R. H. Reaction Mechanism Generator: Automatic construction of chemical kinetic mechanisms. *Comput. Phys. Commun.* **2016**, *203*, 212–225.

(47) Van de Vijver, R.; Zádor, J. KinBot: Automated stationary point search on potential energy surfaces. *Comput. Phys. Commun.* **2020**, *248*, 106947.

(48) Rappoport, D.; Galvin, C. J.; Zubarev, D. Y.; Aspuru-Guzik, A. Complex chemical reaction networks from heuristics-aided quantum chemistry. *J. Chem. Theory Comput.* **2014**, *10*, 897–907.

(49) Bergeler, M.; Simm, G. N.; Proppe, J.; Reiher, M. Heuristics-guided exploration of reaction mechanisms. *J. Chem. Theory Comput.* **2015**, *11*, 5712–5722.

(50) Martínez-Núñez, E. An automated method to find transition states using chemical dynamics simulations. *J. Comput. Chem.* **2015**, *36*, 222–234.

(51) Nett, A. J.; Zhao, W.; Zimmerman, P. M.; Montgomery, J. Highly active nickel catalysts for C–H functionalization identified through analysis of off-cycle intermediates. *J. Am. Chem. Soc.* **2015**, *137*, 7636–7639.

(52) Puripat, M.; Ramozzi, R.; Hatanaka, M.; Parasuk, W.; Parasuk, V.; Morokuma, K. The Biginelli reaction is a urea-catalyzed organocatalytic multicomponent reaction. *J. Org. Chem* **2015**, *80*, 6959–6967.

(53) Ludwig, J. R.; Zimmerman, P. M.; Gianino, J. B.; Schindler, C. S. Iron (III)-catalysed carbonyl–olefin metathesis. *Nature* **2016**, *533*, 374–379.

(54) Dewyer, A. L.; Zimmerman, P. M. Simulated mechanism for palladium-catalyzed, directed γ-arylation of piperidine. *ACS Catal.* **2017**, *7*, 5466–5477.

(55) Jacobson, L. D.; Bochevarov, A. D.; Watson, M. A.; Hughes, T. F.; Rinaldo, D.; Ehrlich, S.; Steinbrecher, T. B.; Vaitheeswaran, S.; Philipp, D. M.; Halls, M. D.; Friesner, R. A. Automated transition state search and its application to diverse types of organic reactions. *J. Chem. Theory Comput.* **2017**, *13*, 5780–5797.

(56) Yang, M.; Zou, J.; Wang, G.; Li, S. Automatic reaction pathway search via combined molecular dynamics and coordinate driving method. *J. Phys. Chem. A* **2017**, *121*, 1351–1361.

(57) Lu, T.; Law, C. K. Toward accommodating realistic fuel chemistry in large-scale computations. *Prog. Energy Combust. Sci.* **2009**, *35*, 192–215.

(58) Van de Vijver, R.; Vandewiele, N. M.; Bhoorasingh, P. L.; Slakman, B. L.; Seyedzadeh Khanshan, F.; Carstensen, H. H.; Reyniers, M. F.; Marin, G. B.; West, R. H.; Van Geem, K. M. 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.

(59) Grimme, S.; Bannwarth, C.; Shushkov, P. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements (Z=1-86). *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.

(60) 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.

(61) Jørgensen, P.; Jensen, H. J. A.; Helgaker, T. A gradient extremal walking algorithm. *Theor. Chim. Acta* **1988**, *73*, 55–65.

(62) Davis, H. L.; Wales, D. J.; Berry, R. S. Exploring potential energy surfaces with transition state calculations. *J. Chem. Phys.* **1990**, *92*, 4308–4319.

(63) Tsai, C. J.; Jordan, K. D. Use of an eigenmode method to locate the stationary points on

the potential energy surfaces of selected argon and water clusters. *J. Phys. Chem.* **1993**, *97*, 11227–11237.

(64) Maeda, S.; Ohno, K. Global mapping of equilibrium and transition structures on potential energy surfaces by the scaled hypersphere search method: applications to ab initio surfaces of formaldehyde and propyne molecules. *J. Phys. Chem. A* **2005**, *109*, 5742–5753.

(65) Maeda, S.; Ohno, K.; Morokuma, K. Systematic exploration of the mechanism of chemical reactions: the global reaction route mapping (GRRM) strategy using the ADDF and AFIR methods. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3683–3701.

(66) Yoneda, Y. A Computer Program Package for the Analysis, Creation, and Estimation of Generalized Reactions—GRACE. I. Generation of Elementary Reaction Network in Radical Reactions—GRACE (I). *Bull. Chem. Soc. Jpn.* **1979**, *52*, 8–14.

(67) Kim, Y.; Kim, J. W.; Kim, Z.; Kim, W. Y. Efficient prediction of reaction paths through molecular graph and reaction network analysis. *Chem. Sci.* **2018**, *9*, 825–835.

(68) Ugi, I.; Bauer, J.; Brandt, J.; Friedrich, J.; Gasteiger, J.; Jochum, C.; Schubert, W. New applications of computers in chemistry. *Angew. Chem., Int. Ed. Engl.* **1979**, *18*, 111–123.

(69) Baker, J.; Kessi, A.; Delley, B. The generation and use of delocalized internal coordinates in geometry optimization. *The Journal of chemical physics* **1996**, *105*, 192–212.

(70) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A.; Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American chemical society* **1992**, *114*, 10024–10035.

(71) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminf.* **2011**, *3*, 33.

(72) Larsen, A.; Mortensen, J.; Blomqvist, J.; Castelli, I.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M.; Hammer, B.; Hargus, C.; Hermes, E. a. The atomic simulation environment—a Python library for working with atoms. *J. Phys.: Condens. Matter* **2017**, *29*, 273002.

(73) Melander, M.; Laasonen, K.; Jonsson, H. Removing external degrees of freedom from transition-state search methods using quaternions. *J. Chem. Theory Comput.* **2015**, *11*, 1055–1062.

(74) Dohm, S.; Bursch, M.; Hansen, A.; Grimme, S. Semiautomated Transition State Localization for Organometallic Complexes with Semiempirical Quantum Chemical Methods. *J. Chem. Theory Comput.* **2020**, *16*, 2002–2012.

(75) Wang, L. P.; Song, C. Geometry optimization made simple with translation and rotation coordinates. *J. Chem. Phys.* **2016**, *144*, 214108.

(76) Aldaz, C.; Kammeraad, J. A.; Zimmerman, P. M. Discovery of conical intersection mediated photochemistry with growing string methods. *Phys. Chem. Chem. Phys.* **2018**, *20*, 27394–27405.

(77) Frisch, M. J. et al. Gaussian 16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

(78) Maeda, S.; Harabuchi, Y. On benchmarking of automated methods for performing exhaustive reaction path search. *J. Chem. Theory Comput.* **2019**, *15*, 2111–2115.

(79) Susnow, R. G.; Dean, A. M.; Green, W. H.; Peczak, P.; Broadbelt, L. J. Rate-based construction of kinetic models for complex systems. *J. Phys. Chem. A* **1997**, *101*, 3731–3740.

(80) Pfaendtner, J.; Broadbelt, L. J. Mechanistic modeling of lubricant degradation. 2. The autoxidation of decane and octane. *Ind. Eng. Chem. Res.* **2008**, *47*, 2897–2904.