ANNUAL REVIEWS

The Exploration of Chemical Reaction Networks

Jan P. Unsleber and Markus Reiher

Laboratory for Physical Chemistry, ETH Zurich, 8093 Zurich, Switzerland;
email: markus.reiher@phys.chem.ethz.ch

Keywords

Abstract

Modern computational chemistry has reached a stage at which massive exploration into chemical reaction space with unprecedented resolution with respect to the number of potentially relevant molecular structures has become possible. Various algorithmic advances have shown that such structural screenings must and can be automated and routinely carried out. This will replace the standard approach of manually studying a selected and restricted number of molecular structures for a chemical mechanism. The complexity of the task has led to many different approaches. However, all of them address the same general target, namely to produce a complete atomistic picture of the kinetics of a chemical process. It is the purpose of this overview to categorize the problems that should be targeted and to identify the principal components and challenges of automated exploration machines so that the various existing approaches and future developments can be compared based on well-defined conceptual principles.

1. INTRODUCTION

Any chemical process may be decomposed in terms of a network of elementary steps. The exact knowledge of all elementary steps, including intermediates, transition structures, and products allows for kinetic modeling and the prediction of concentration fluxes through the network. Due to advances in the field of theoretical and computational chemistry in the past decades and the ever-increasing computational power of modern hardware, it has become feasible to explore chemistry on a broad scale, i.e., tackling the vast dimension of chemical reaction space. In general, the truly exhaustive exploration of some chemical process cannot be guaranteed, but the current state of theoretical and computational chemistry allows for the generation of algorithms that march into an enormously larger fraction of this space than what would be accessible by manual exploration.

Chemical reaction space exploration comprises methods that generate knowledge about a chemical mechanism through atomistic modeling. These methods can be closely tied to data-driven approaches, which attempt to induce reliable results about potential reactivity from existing experimental knowledge in the chemical literature (1). Reactive pathways are then inferred from rules and models generated on the basis of existing data (2–4), such as reactions published in patents (5, 6) or curated databases (7). These approaches were initially devised to exploit expert generated rules (8, 9) and have recently seen an impressive revival owing to the rise of machine learning techniques (5–7, 10–16). Machine learning facilitates automating and abstracting the rule generation based on amounts of data that were unprocessable before. This data-driven approach can be a powerful means to generate ideas about chemical reactivity for new systems similar (according to some measure) to reactants studied in the literature. By contrast, chemical reaction mechanism exploration based on quantum-chemical first-principles methods, which we consider in this work, provides an option, next to experimental synthesis, for probing the validity of the ideas derived from data-driven inference, provided that depth, reliability, and accuracy of the quantum-chemical exploration can be guaranteed. One aim of this overview is to work out the criteria that need to be considered and fulfilled in order to make quantum-chemical mechanism exploration a reliable peer-to-data-driven reactivity deduction.

In a recent review (17), we discussed in detail algorithms for chemical reaction space exploration. They may be grouped into three classes (17): (*a*) those that aim at a complete exploration of a given potential energy surface, (*b*) those that trade breadth for depth by relying on structure hopping and chemical heuristics, and finally, (*c*) those that exploit human intuition to tame the combinatorial explosion of structures in vast networks involving numerous possible reactants and pathways through, e.g., steering by interactive quantum mechanics. As further reviews on the topic can be found in References 18 and 19, we only mention some key methodological work in the field, highlighting that a variety of algorithms and concepts have already been devised. They comprise graph-based approaches (20–22), first-principles heuristics that extract rules from the conceptual interpretation of the electronic wave function (23–25), chemical heuristics using first-principles calculations (26, 27), and stochastic approaches (28, 29). It is not surprising that the global extension of local search and sampling methods has delivered an even broader range of exploration algorithms. Examples are those that exploit artificial forces (30–33), growing string methods (34, 35), exploratory ab initio molecular dynamics (36, 37) metadynamics (38–41), and other enhanced sampling methods (42–45).

Considering that exhaustive overviews on the topic have already been provided (17–19), this work focuses on comparability and the missing links needed to make the different approaches comparable. Because of the numerous concepts and algorithmic procedures introduced so far, it is difficult to have a balanced comparison of the different exploration schemes on the same footing. This becomes even more difficult when considering that exploration algorithms may be designed

to serve specific purposes (e.g., gas-phase versus solution chemistry, restrictions to specific compound classes or computational methods, and so forth).

There have been some attempts in the literature to compare existing algorithms at the example of specific target problems (46, 47). However, such comparisons are difficult. A trivial hurdle turns out to be the parameters and thresholds that control an exploration algorithm, which may be chosen in a nonoptimal way, making a direct comparison less conclusive (47). Furthermore, the performance of some algorithm may also depend on the choice of a specific task. For this reason, it is desirable to have a set of criteria at hand, which allow one to arrive at informed and balanced conclusions about a specific computational approach.

For this reason, we adopt a metaconceptual perspective aimed at the definition of common general concepts and requirements. First, we discuss and categorize the range of possible mechanistic targets. To keep this task well defined and controllable, we focus on the identification of elementary reaction steps and hardly touch upon the natural extension and combination with subsequent kinetic modeling. For work relating to kinetic modeling in the context of reaction networks, we point the interested reader to References 48–64.

2. CATEGORIZING MECHANISTIC SEARCHES

The central paradigm of reaction mechanism exploration is the idea that a chemical process under consideration can be mapped onto a (transformation) network of elementary reaction steps connecting reactants and stable intermediates through transition-state structures. As a starting point, we may adopt the simplest concepts as they emerge from Eyring's absolute rate theory (65) or Kramers's theory (66, 67). For the sake of clarity, this may be considered sufficiently fundamental as a network of elementary reaction steps can be arbitrarily refined through subsequent calculations that provide data required for more advanced rate theories.
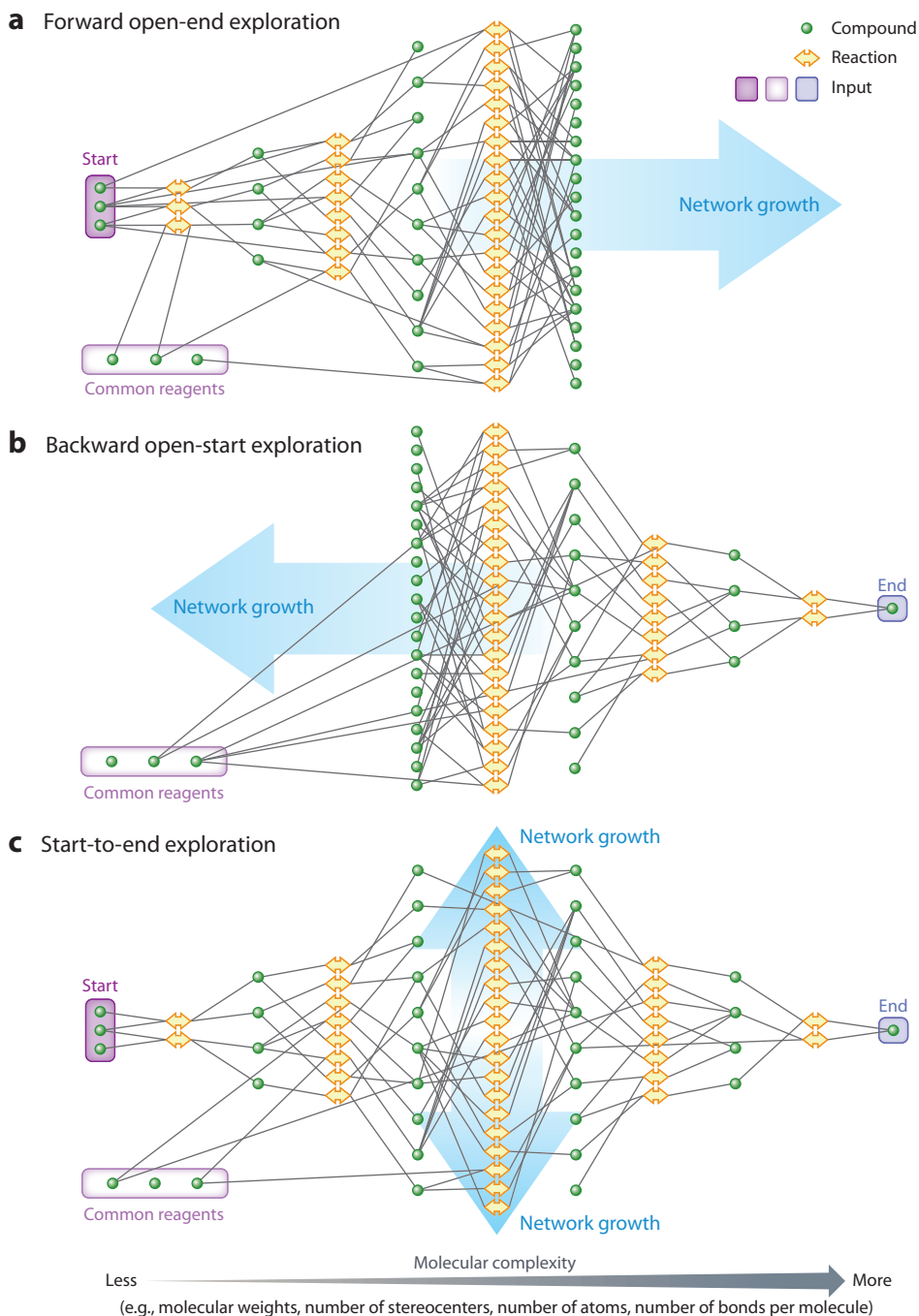
We consider all parts of chemical (reaction) space relevant for a problem as one huge network that encodes a chemical function. Because an exploration process will have some starting point and might already target a specific end point, we define three principal exploration types: a forward exploration with an open end (FOE), a backward exploration with an open start (BOS), and a start-to-end exploration (STE) with known start and end (see **Figure 1** for a graphical representation of these exploration types).

Asking how compound A reacts with compound B manifests a prototypical FOE example with a potentially open end. The compounds A and B constitute a known start. We must then assume that any stable intermediate of the detailed mechanism may, in principle, again react with the starting compounds A and B, or with any other intermediate provided that it is sufficiently long-lived. By contrast, the task of how to synthesize compound C constitutes an inverse problem (68–70), which is difficult to solve as the number and type of reagents in forward direction is not known in the beginning. Hence, this BOS exploration type will in general require additional information on potential reagents and catalysts that can promote the forward reaction. The machine learning–based retrosynthesis algorithms mentioned in the Introduction can provide valuable hints on which reagents to choose. If starting and end points of an exploration are for some reason known and fixed, an exploration algorithm may identify the viable network of connecting elementary steps and then contribute to the question of what is the (according to some measure) best pathway and how to promote it.

Because the final product(s) of a chemical process may take part in subsequent reactions and because chemical reactions are in general reversible, the labels "start" and "end" can become somewhat arbitrary. Still, the definitions of the three types discussed so far represent typical classes of exploration problems.

Three principal exploration types: (*a*) forward open-end exploration, (*b*) backward open-start exploration, (*c*) start-to-end exploration. Starting material (initial reactants) are given on the left-hand side and are typically of small molecular weight. Common reagents (including ubiquitous molecules in the environment) are depicted separately to highlight their multipurpose application. A network can grow exponentially in forward explorations, because all newly emerging stable intermediates or side products need to be considered as new reactants.

a Forward open-end exploration

Start

Common reagents

Network growth

Compound
Reaction
Input

b Backward open-start exploration

Network growth

Common reagents

End

c Start-to-end exploration

Network growth

Start

Common reagents

End

Network growth

Molecular complexity
Less ————————————————————— More
(e.g., molecular weights, number of stereocenters, number of atoms, number of bonds per molecule)

Although all reagents that are part of the network are treated in the same way, it can be advantageous to single out specific compounds due to their ubiquitous or universal nature as reactants such as, for example, H_2O, $^1O_2$, small elimination products (e.g., HBr), (solvated) protons, and so on. In this way, screening tasks such as those for a reaction's robustness toward functionalized molecules (71) can be simplified. Further conceptual steps, such as work-up, reactant separation, and sequential reaction steps, can be easily encoded into a network by restricting the connectivity of certain nodes. Note, however, that a specific reference to an actual environment of a reaction is not made in the labels introduced so far. Implicit and explicit solvation, as well as metal surfaces, protein environments, or other environments will initially lead to different, though related, networks.

The setting defined so far comprises a broad range of scenarios from simple reaction mechanisms in the gas phase (STE without additional common reagents) to autocatalyzed reactions with side reactions (FOE with additional common reagents), to probing a restrosynthetic proposal for synthesis (STE with various restrictions and constraints), to screening of an optimum catalyst for a specific target transformation (STE with different catalysts as common reagents). Classifying reaction network exploration protocols according to the three types introduced above allows us to categorize the different algorithmic and conceptual approaches toward the exploration challenge and compare them on the same basis. Considering the open nature of FOE-type explorations, it is obvious that an actual algorithm will require some adjustable determination of the termination of the exploration. For the termination of FOE and BOS explorations, bounds on molecular weights are simple criteria, and in the case of STE explorations the number of reactions between start and end could be limited. Furthermore, constraints based on kinetic modeling at a certain temperature and other external constraints can be introduced.

3. GENERAL NOMENCLATURE FOR REACTION NETWORKS

Because notions such as system, molecule, compound, reagent, and structure are often used interchangeably, we introduce and define the essential concepts. To define the components of a reaction network, we first settle on the notion of a molecular structure.

Molecular structure: A single arrangement of atomic nuclei on a Born–Oppenheimer potential energy surface shall be denoted a molecular structure. Consequently, it is represented by a set of Cartesian coordinates for the nuclei. Depending on the interpretation in terms of a kinetic theory, we are, in general, interested in structures with specific properties, most importantly those with vanishing geometry gradients that highlight stable intermediates and first-order transition states.

Properties: Each such molecular structure may then be assigned some property that relates to an electronic structure defined by the number of electrons and spin assigned, its state of excitation (typically the ground state), and any derived molecular property (such as the electronic energy or total dipole moment). Here, we may also allow for conceptual properties such as partial charges and local spin that rely on some chosen decomposition scheme.

With these basic terms, we may proceed and introduce a compound:

Compound: A set of molecular structures with the same nuclear composition and connectivity in terms of chemical bonds denotes a chemical compound. The bonding pattern that is crucial for the assignment to the set may be either fixed based on heuristic rules or, preferably, determined through a bonding analysis [e.g, in terms of Mayer bond orders (72)].

Although such a bonding analysis may often be unique (and particularly helpful to detect and assign molecular structures as different conformers to the same compound), it cannot be excluded that certain structures end up in a gray zone (for instance, if bond orders are very low). This may then create ambiguities on the conceptual level of compounds, but not on the more fundamental level of the molecular structures (characterized by a vanishing gradient).

Whereas we consider structures essentially as static for quantum-chemical single-point energy calculations, the concept can be extended to molecular dynamics simulations. Then, however, clustering algorithms (73–77) are required to assign sets of structures to a node, which is then a generalization of the compound concept. Transitions between such compounds derived from clustering trajectory data can provide access to kinetic information through Markov state modeling (51, 78–80). A single-point-calculation approach requires the explicit generation of all relevant conformers (for example, with the program package RDKit),[1] whereas molecular dynamics simulations then generate compounds through clustering.

In order to generate relationships between structures and eventually compounds, transformations from one node in the network to another one must be uncovered. In general, one first strives to identify all relevant elementary steps that connect molecular structures and then later generates a reaction network out of these elementary steps that transform compounds into one another. For transformations we use the following terms:

Elementary (reaction) step: One (or more) structures may be related to one (or more) different structures via a single transition state representing an elementary step that can be resolved as a sequence of structures along a corresponding minimum energy pathway.

Reaction: A chemical transformation of one (or more) compounds into one (or more) different compounds represents a reaction that can be distilled from at least one, but in general more than one, elementary steps.

As an example, **Figure 2** shows a small reaction network built from the four types of nodes defined above.
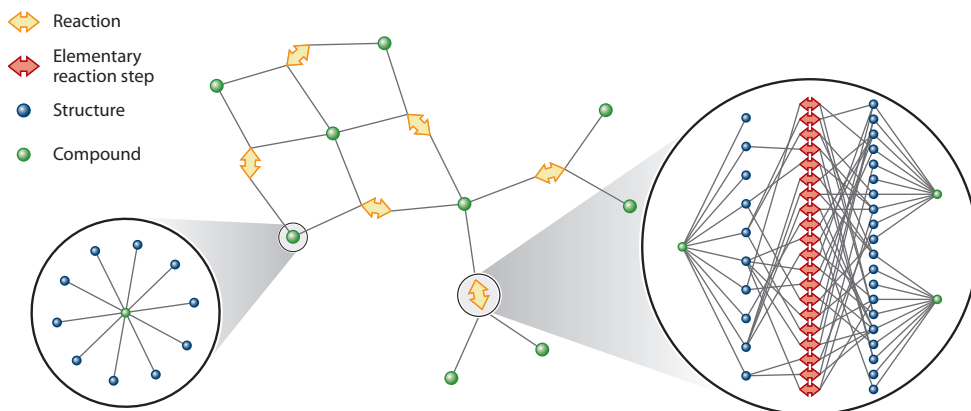

Figure 2

Example network built from reactions and compounds. The expanded section shows a subnetwork structure that highlights the relationship between structures and compounds as well as the relationship between reactions and elementary steps.

---

[1] RDKit is an open-source cheminformatics software, available at **http://www.rdkit.org**.

Any algorithm tackling a chemical question posed in terms of changing molecular structures should explore the transformation steps of these structures and eventually generate compounds and reactions. Hence, chemical reactivity and function will then be encoded in terms of the elementary steps connecting structures as nodes of a network. However, we may introduce a few more general concepts that allow for comparison between compounds and reactions. These generalizations are important to cover concepts such as residue that assign a spectator role to certain parts of a molecule; they are indispensable in an explicit quantum-chemical description, which, in general, requires detailed knowledge about all electrons and nuclei involved. As such, these generalizations represent the first step toward mechanistic interpretation, because the network of structures and their connecting elementary steps are agnostic with respect to the assignment of any role in the chemical function under consideration. Furthermore, generalizations can be invoked to assign networks that are very different, on the level of atoms and elementary particles, to the same class. Typical examples are generalizations of molecular function to be fulfilled by a comparatively wide range of compounds, such as oxidizing agent or hydride source. For these reasons, we introduce three additional terms:

State of a compound: Reducing or oxidizing a compound (and hence, all its structures), changing the degree of excitation or the spin state, or protonating one of its active sites leads to different molecular structures in a quantum-chemical description and would therefore produce different nodes in a reaction network from which new elementary steps lead to further structures. The close relationship to quantum mechanical concepts and prototypical reactions (protonation, electron transfer) make them amenable to automated identification and assignment in quantum-chemical procedures.

Reagent: A group of compounds that share the same role in a network form the category of a reagent that is specific for that role. Typically, reagents are "oxidant," "reductant," "acid," "base," and "catalyst."

Purpose: The concept of "reagent" allows for comparison and a higher degree of abstraction of a network, which for the network itself is its "purpose." Purpose is the superordinate concept to "reaction." A typical purpose in reaction chemistry is "synthesis" or "catalysis," but the term may as well refer to any molecular functions encoded in terms of structural changes such as "molecular motor" or "mechanochemical device."

To assign an oxidation, charge, spin, excitation, or protonation state to a compound not only attributes chemical meaning to a network of structures but also has consequences for network data handling, storage, analysis, and presentation. The fundamental nature of these states also allows for their monitoring during exploration (81). Furthermore, note that some compounds may be part of different sets of reagents in different contexts. Early stage exploration algorithms do not feature these last two types of nodes and how they encode purpose, because of the bottom-up nature of reaction network exploration based on the first principles of quantum mechanics.

Given two or more explored networks, these concepts define connections between them, which can and should be exploited by exploration algorithms. One type of connection is node sharing, i.e., the fact that a compound (and hence, its structures) can occur also in another network. Calculated data on such parts of a network can be reused and call for centralized database storage (the `QCArchive` project by MolSSI[2] is one example of a technical realization that could be exploited by network exploration algorithms). Another type of connection is contextually joining reaction networks of structures that are very different on a quantum-chemical level through the assignment of

---

[2]The `QCArchive` database is available at **https://qcarchive.molssi.org/**.

the same purpose. A typical example is a generic catalytic cycle, which on a realization level can put networks with different catalysts or with differently substituted substrates into the same context. As a result, one could even derive functional networks in which modulating effects of molecular structure are hidden in an environment description of a quantum-chemical embedding approach.

Deliberately, we have left some terms unspecified to allow for some vagueness in the context of network exploration. This is convenient as it facilitates comparability of different algorithms and does not interfere with the core concepts introduced above. For instance, we have not specified the term "system" and may use it as a stand-in for the way possible elementary steps are actually set up and then explored in an algorithm.

4. FEATURES OF REACTION NETWORK EXPLORATION ALGORITHMS

4.1. General Foci

The main results of a reaction network exploration will be extracted at the level of compounds and reactions. Context-based abstractions into reagents and purpose then allow for arranging and highlighting the vast amounts of data. Splitting up reactions and compounds into structures and elementary steps allows for understanding the fine details of a particular mechanism at the level of the physical dynamics of the involved molecular species.

At the level of compounds and reactions, it is reasonable to distinguish between two dimensions: the breadth of a network and its depth. Here, breadth refers to the amount of reactions and compounds that have been incorporated into the network, whereas depth denotes the amount of structures and elementary steps discovered for each of the compounds and reactions. Naturally, both will be difficult to determine in absolute terms as this would require complete knowledge about a chemical process, which is the target of an exploration algorithm.

A lack of depth will likely yield qualitatively wrong kinetics, and hence, it will predict wrong distributions of reactants across the networks once these are modeled. Depth fidelity also comprises the accuracy of the chosen computational methods, which manifests itself in errors on barrier heights. A lack of network breadth will yield qualitatively wrong results for the total kinetics as it implies that important side reactions have been missed.

In order to describe a chemical process as reliably as possible, both dimensions must be pushed to their limits, and it will be important to find algorithm-intrinsic measures that hint toward their saturation. Therefore, it is instructive to define two distinct types of completeness, one for the breadth, the graph fidelity, and one for the depth, the node fidelity.

Naturally, different applications can require a different focus with respect to these directions. Consider the following two examples:

- The calculation of the feasibility of a single reaction cascade to gain insights into its reaction mechanism by exploring a multitude of conformers and pathways requires slow growth of the network breadth and very accurate exploration of its depth.
- The exploration of an entirely unknown set of reactions of given compounds may favor a quick growth, with an a posteriori automatic refinement scheme to increase depth and breadth for interesting parts of the network. A low accuracy in terms of depth for unaccessible and therefore uninteresting parts of the network will be sufficient and often mandatory for the sake of feasibility.

As should be apparent now, a multitude of requirements exist for a generally applicable exploration tool. It shall be able to switch between the different exploration modes defined in Section 2

and arrive at a sufficient accuracy. To this end, we define a more complete set of key challenges for exploration algorithms in the next section that allows one to classify and assess algorithmic developments regarding their scope and capabilities.

4.2. Challenges

In order to later compare and eventually rank algorithms that explore chemical space, computationally it is imperative to define their goals. These targets presented in the next subsections are driven by the following general challenges:

Validation challenge: The exquisite details that exploration algorithms can generate for any chemical process raise the question of reliability as, in general, no or very little experimental or theoretical reference data will be available. Consequently, uncertainty quantification (82) will become a crucial part of the whole exploration process (63, 83).

Operating with huge amounts of raw data: Automated exploration algorithms will be most useful for cases that require thousands or millions of structure searches and optimizations. As a consequence, huge amounts of data will be produced whose manual inspection is not at all possible. Therefore, automated exploration algorithms must be very stable, handle all data in a fully automated and integrated way, and automatically draw the operator's attention to critical situations such as convergence failures that cannot be resolved automatically.

Minimal expectations on the operator side: The software design that implements an exploration algorithm must consider that all its parameters, thresholds, screws, and bolts cannot be fully understood on the application side. In other words, the efficiency and reliability of the algorithm should depend as little as possible on knowledge about the intricate effects of changing some of its parameters. Naturally, the default values for them should be expected to be stable and applicable to a wide range of problems and situations.

Unknown degree of incompleteness of generated data: For any reasonably complicated case, it cannot be rigorously proven that the exploration of the corresponding reaction network of interest has been completely explored. Therefore, it will be difficult to construct sufficiently complex benchmark cases against which exploration depth and breadth can be measured, and most likely, these will only emerge in a joint effort of various approaches over time.

4.3. Targets

Within this basic setting, we may now formulate goals and targets of exploration algorithms. Some of these goals have been tackled already, but all of them remain to be improved upon in some way. It is therefore decisive to have this list at hand for future developments in the field.

4.3.1. General applicability and stability.

A key goal for the exploration of reaction space must be flexibility with respect to the class of molecules and their environments that can be considered. An algorithmic restriction toward a specific compound class, reaction environment, or state of aggregation can be a severe limitation in the discovery process as it would severely limit the domain of exploration and exclude potentially decisive reagents, solvents, and so on. An ultimately useful protocol must be able to accommodate any potentially relevant molecular scenario, even if it is truly hard to map in a virtual screening process. Examples can be found in structures with multiple transition metals in a protein environment or with molecules on metal surfaces.

Furthermore, it must be ensured that the algorithms employed are as stable as possible because an exploration protocol may lead them to regions of configuration space that pose problems for

them. For instance, orbital convergence will be a key issue—solution methods for complete failure of convergence (84) or for convergence to wrong solutions (85) are required.

It will, in general, be necessary to operate with a wide range of approaches, spanning fast and less accurate as well as slow and accurate energy assignment protocols, from detailed structure construction and search to advanced sampling through molecular dynamics and Monte Carlo methods. Therefore, the requirements for exploration algorithms in terms of general applicability are immense.

4.3.2. Intrinsic constraint monitoring and adaptation of the exploration algorithm.
Obviously, any algorithm will involve choices that eventually limit its applicability. For instance, exploring reaction space on the basis of individual evaluations of the stationary Schrödinger equation emphasizes the role of the electronic energy, which will be reasonable as long as it determines the major part of the relevant energy differences. When the exploration then enters a regime of a potential energy surface that is rugged, this needs to be substituted by a proper molecular dynamics or Monte Carlo sampling approach—and vice versa. The quantum-chemically explored reaction network can then be supplemented with proper information from some enhanced sampling approach, possibly encoded in terms of a kinetic model built from this sampling [e.g., through Markov state models (51, 78–80)].

In order to recognize these limitations, it is crucial to determine the limits with algorithm-intrinsic means while exploring a network in order to achieve sufficient breadth and depth. In the example above, a large number of small barrier heights in some region of reaction space can be taken as a sufficient indication. As long as no fully satisfactory integrated software is available, interoperability of different implementations will be of key importance.

4.3.3. Taming conformational explosion.
With increasing size of molecular structures, conformations of a compound become increasingly important. For moderately sized compounds, conformers can be explicitly constructed (24, 86) and optimized. At some point, however, this will no longer be feasible, and again, sampling approaches will be needed. For truly large molecules, this will eventually become a cumbersome task as highlighted by the protein folding problem. The whole menace of conformational depth becomes even more severe when considering transition state structures, especially when embedded into a fluxional environment such as, for instance, water. An exploration based on a single conformation per compound could be a first step toward generating an overview on the most interesting parts of a network of elementary steps, but later refinement will then be more mandatory.

4.3.4. Type of energy data provided.
It is obvious that ultimate energetical data assigned to nodes of a network should be free-energy data within a well-defined thermodynamic ensemble. However, this ultimate goal is not easily accessible to arbitrary accuracy, and a first step would therefore be a network based on electronic energies only (to be later refined by modeling contributions from the nuclear framework at some finite temperature). Given the fact that macroscopic constraints such as temperature, pressure, volume, particle number, and so forth can change, it would be desirable to store all data per structure from calculations, which allow one to evaluate free-energy data for changing external parameters at will. Arriving at free energies may be accomplished within very different models ranging from those that start with the standard rigid-rotor harmonic-oscillator particle-in-a-box model (see, e.g., Reference 57), to continuum solvation approaches (see, e.g., Reference 87), to additivity schemes (see, e.g., Reference 88), and to explicit sampling approaches (see, e.g., References 89 and 90).

4.3.5. Environment embedding.
Exploring some chemical function or reactivity is most easily accomplished for reactants lacking any environment (i.e., for isolated species that may be considered to represent a gas-phase situation at low pressure). The inclusion of a suitable environment represents a major challenge that must be addressed by an exploration algorithm. Clearly, explicit molecular dynamics in a box of sufficient size under periodic boundary conditions is an attractive choice, but this quickly limits the scope of an exploration to the Born–Oppenheimer surface that can be constructed for the elementary particles in that specific box. Changing reagents will not easily be possible. Hence, embedding schemes (for reviews, see References 91–94) play an important role, possibly combined in a multilayer strategy that extends from explicit environment structures close to a reactant to structureless dielectric environments at large distance. Different embedding scenarios will lead to variations of similar networks that require specific care on the data-management size in order to allow for relevant chemical interpretation and conclusion.

4.3.6. Error and uncertainty diagnostics.
Any model that produces raw data for a network exploration algorithm will rely on certain approximations. For instance, electronic structure models are likely to be based on fast semiempirical (95) or density-functional theory (96) methods, which can be affected by surprisingly large errors for specific structures (see References 97–99 for examples). Without proper uncertainty quantification, explorations based on such data will be inherently unreliable. Error estimation (82) therefore becomes key. Methods have been devised that point the way for how this can be achieved (57, 83, 100–102).

4.3.7. Automated error reduction.
Naturally, nonnegligible errors for certain structures detected during exploration require (automated) refinement by launching more accurate energy calculations. Typically, this will encompass starting reliable ab initio calculations. Although this is feasible for single-reference methods such as explicitly correlated coupled cluster theories (103), which can typically be run as black boxes, this is not that straightforward for multiconfigurational problems although significant progress has been made in this respect (104–106). Then, only feasibility considerations are an issue, but may be alleviated by embedding calculations (94, 107–110). If an exploration protocol allows for such calculations on demand, then the determined error should be propagated through the network in order to reduce the uncertainties for the approximate exploration method. As we have shown (83, 111), this is possible by means of machine learning. Such approaches are determined to play an important role in maximizing the accuracy while minimizing the computational cost of the network explorations.

4.3.8. Intuitive and immersive interaction and visualization.
As vast networks will contain too much data for a human to grasp, it is imperative that results of any exploration can be displayed in an accessible way and that conclusions can be drawn with the help of algorithms. This requires specific software for human–machine interaction as visual inspection of alphanumerical data will be unfeasible and pointless. Suitable graphical user interfaces are required that also free one from looking at alphanumerical raw data stored in some data bank container. They may invoke new hardware (see, for instance, References 112–116) to intuitively experience these data in order to easily put focus on relevant aspects. Such hardware and software also allow for manual interference and control of the otherwise automated exploration process.

Obviously, interaction cannot mean manual editing of input files for a specific program and a manual database query and insertion, but rather it means a simple click on a node, opening a context-dependent minimal menu. In view of the definitions introduced above, it is possible to define layers for this purpose reflecting in the first layer structures and elementary steps, and in the second compounds and reactions, and then on top of this one abstract layer that encodes reagents

and purposes. **Figure 2** shows one example of such a representation, also indicating the idea of switching between these layers.

It appears natural that interacting with such an interface and steering the whole exploration protocol would be easier (and for nonexperts most convenient) through language processing algorithms. Voice control of exploration is one option that has already been proposed as a valuable goal for computational chemistry (117, 118). Given the advances in the context of daily-life artificial-intelligence assistants and the availability of open-source language processing libraries (119),[3] this task is not at all difficult and even the setup of an interpretation assistant, which can translate between chemical jargon and the precise ingredients required for quantum-chemical calculations on a properly prepared atomistic model, will be rather straightforward.

4.3.9. Maximum accessibility. Although many quantum chemistry programs are reasonably easy to install, an exploration software will most likely involve a database, a user front-end, and a back-end that manages all calculations. Hence, installation and therefore accessibility are more complex by design. Both hardware and quantum-chemical software packages available at the outset of an exploration attempt will result in various setups with different computational capabilities. Given the recent rise of cloud computing and virtual machines, easy access to standardized configurations of the back-end accessible as images is likely a key feature that allows easy setup of the entire machinery for nonexperts. Although proprietary software can be made accessible in this way, open-source initiatives are likely to form a more stable and user supported base. The latter are preferred in any case due to better reproducibility, reliability, and fidelity. With the plethora of settings already available for quantum-chemical calculations and the profound impact they have on the resulting data (120), it will only be possible to reproduce and understand data if their production and processing can be inspected at the source-code level. Given the additional algorithmic layers needed in a reaction network exploration it is of key importance to document and explain possible settings. For the same reasons, a comprehensive way of accessing and documenting the used options for a given exploration run is another key part of this goal.

4.3.10. Data transferability. For efficient explorations it is highly desirable that already generated networks and the contained data can be reused and combined such that subsequent explorations do not need to be recalculated but instead can be seamlessly incorporated with existing data. Considering the vast amount of standard reactions and the resulting list of frequently used reagents and often occurring reactants, data integration will be highly efficient. However, integration of existing data will only leverage an exploration attempt if uncertainty quantification has rated and labeled the existing data. It is therefore desirable to generate a central library or database of highly accurate results that is continuously extended with data from local explorations, if that data has been generated with an accuracy above a certain threshold. A central database of chemical reaction space would also be a promising starting point for the application of meta-algorithms and machine learning models to exploit and learn from the reaction chemistry that has been mapped out.

4.3.11. Enhanced kinetic modeling. Eventually, kinetic modeling will be required to study concentration fluxes through a network. As it is a priori not clear what the number and kind of elementary steps will be that constitute the network, rather general microkinetic solvers will be required, possibly tightly entangled with the exploration algorithm itself to enhance and guide the latter (63). Consequently, such solvers should be capable of dealing with vast timescale ranges and

---

[3]Mycroft (**https://mycroft.ai/**) and Jarvis (**https://openjarvis.com/**) are examples of open-source voice-based control technology.

varying degrees of molecularity (at least up to second-order kinetics). Whereas special tools (121) and commercial solutions (122) are already available, it is obvious that reaction network exploration presents further challenges that demand more developments. Furthermore, to accurately model the flux in reaction networks, theories beyond Eyring's absolute rate theory will be needed (123–127). Clearly, quantum tunneling must also be considered (128). However, improvements on this level will represent a natural extension of a network of elementary steps, as the software may automatically gather more information about the potential energy surface in the vicinity of the nodes in such a way that advanced rate theoretical expressions can be evaluated.

5. OPTIONS FOR COMPARISONS OF NETWORK EXPLORATION ALGORITHMS

Comparisons of existing algorithms that explore reaction networks have already been attempted (46, 47). However, in view of the many aspects needing consideration in such an attempt and the fact that a balanced set of benchmark reaction networks resembling various scenarios of practical relevance would be required, such comparisons will not be easy, although they are certainly needed. It is important to assemble a set of criteria by which an exploration algorithm should be assessed and rated in order to highlight the different problem classes that need to be addressed. It will be necessary to create a multidimensional diagnostic to characterize a specific exploration protocol.

Ultimately, any such measure will have to be based on solid comparable data, and hence, benchmark data should be generated and collected in future work. Given the many different types of chemical processes that could be probed, an extendable set of smaller networks appears to be the best choice in practice. Using the three initially defined exploration modes (FOE, BOS, and STE; see Section 2) we may propose that three main tests be run for each benchmark network. Because automated network exploration has a significant and important technical and implementational component, extensibility and access should be measured. Furthermore, the reproducibility in terms of the licensing of the source code should also be a factor. Assuming a scale from one to ten, a software that is closed source and requires a specific set of commercial software without any possible extension would be ranked zero. A fully open-source software with the possibility of interfacing with any other quantum-chemical code and a set of open-source default programs provided with the exploration tool to guarantee full functionality would rank as ten.

For a fine-grained comparison of exploration algorithms and implementations assessed at benchmark networks, five key descriptors (sketched in **Figure 3**) should be sufficient. The five descriptors first and foremost include the two types of reliability measures described above. The node fidelity (depth) is calculated by comparison of the energy spectra of the different nodes with the reference data. The network fidelity (breadth) is calculated as the amount of correctly identified compounds and reactions. Naturally, the expense of the calculations is also one of the descriptors. The computational cost could be measured as done in References 46 and 47, which use the number of gradient calculations as dominant measure. Because the accuracy depends on the methods used and that gradient calculations with different methods can vary significantly in duration, the measure should also incorporate an average time of a gradient calculation on a certain hardware. In order to probe the transferability of networks, meaning the possibility of reusing previously generated network data, we may measure the acceleration of an exploration when starting from a subset of already existing nodes in the reference network and may also supplement the running exploration on the fly with structures surfacing during the exploration that are found to be already in a central database. Furthermore, nodes that are automatically determined to be similar (e.g., nodes that were calculated with different solvents as environment) can and should be
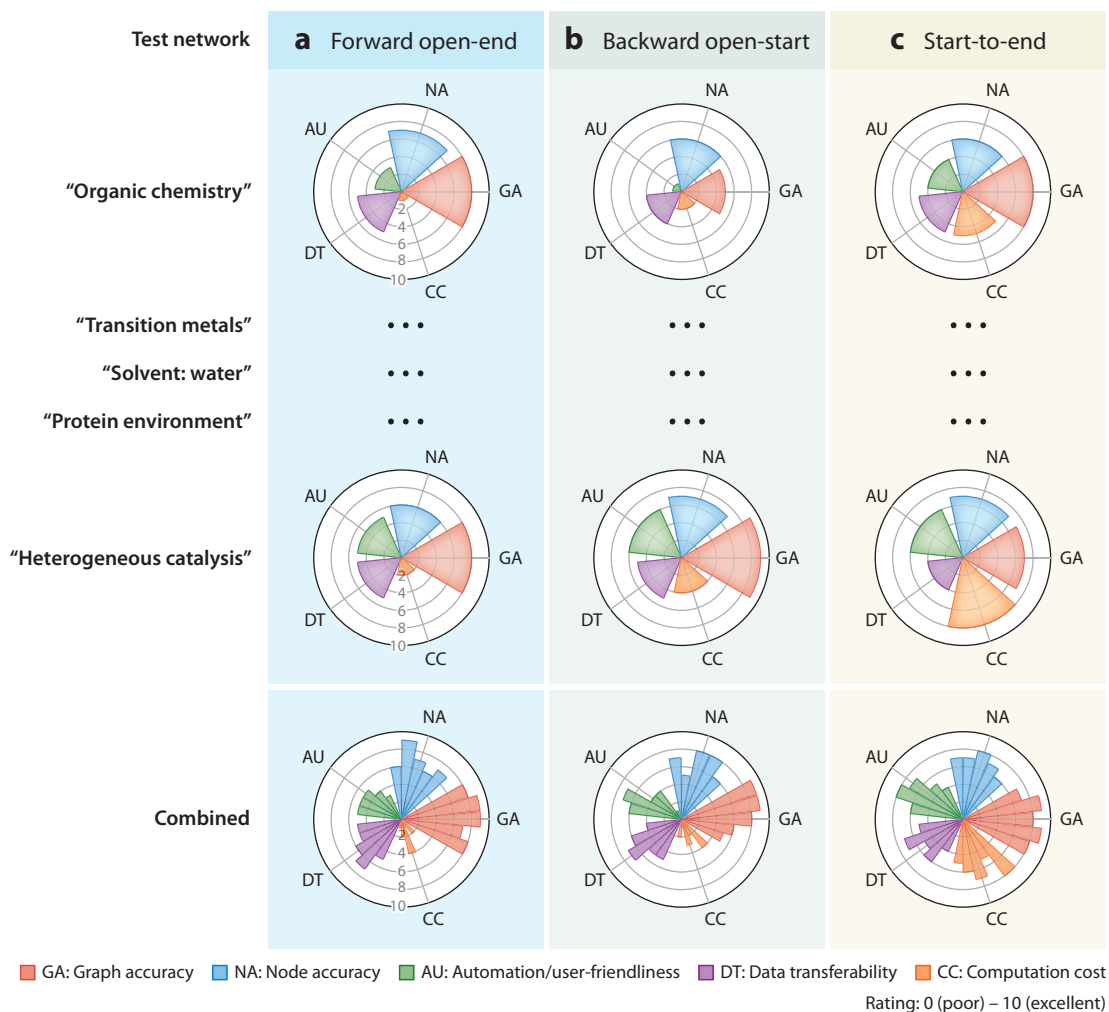
Figure 3

Fictitious benchmark data shown in rose plots for chemical reaction network exploration software.

used as starting guesses. In order to credit the fact that these software packages and algorithms should eventually be useful to nonexperts in the field of computational and theoretical chemistry, it is important to measure the user-friendliness (immersion) as well as the degree of full automation of a given protocol. The amount of human work required to steer and control an automated network exploration software could be taken as a measure for user-friendliness and for the level of automation achieved. Unfortunately, this may require extended user studies. Given the vastly different approaches to solving many of the problems discussed in this work, a challenge similar to the Cambridge Crystallographic Data Centre organic crystal structure prediction challenge (129) or the Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) challenge (130) could be a desirable testing ground for the field. It would allow for better comparisons for subtopics as well as help focus the efforts on particular unsolved problems.

6. ASSESSMENT OF THE CURRENT STATUS

So far, we have discussed general concepts and goals for reaction network exploration algorithms and their implementation. To assess the current status of the field we highlight some aspects according to the literature. First, it is noteworthy that most exploration methods employ existing quantum-chemical software packages and models. They generally do not implement new electronic structure models in a closed-source form, which has advantages for availability, reliability, and reproducibility of the overall method. The rigorous estimation of the associated errors on structures and energies, their propagation, and minimization during the exploration has been tackled only in a few instances (57, 58, 82, 83). In terms of the scope of the chemistry that has been the subject of exploration attempts so far, most methods work routinely for basic (gas-phase) organic chemistry, and some of the reported algorithms have been shown to work with transition metals (18, 20, 21, 23, 25, 29, 40, 43, 131). The explorations reported so far have largely avoided explicit description of environments such as solvents or protein embedding. Most applications reported for exploration algorithms have been of the proof-of-principle type. It is therefore still a long way ahead of us until we have general implementations available that can be routinely employed in daily research. No study has reported data on a network with, say, 1,000 confirmed compounds and 10,000 possible reactions. (Note that this implies $\gg 1,000$ conformers are to be generated and $\gg 10,000$ elementary steps are to be calculated.) The challenges of the visualization and analysis of such an amount of data and the transferability of subspaces of the network have only been touched upon (23, 24). Overall, however, important developments have been accomplished and key steps have been taken (see Section 1). Considering three major requirements, i.e., a robust yet flexible exploration algorithm, a scalable and extendable back-end for the actual computations, and kinetic modeling relating the generated data to actual chemistry, we note that no single software framework has been established so far that accomplishes these goals so that comparisons of implementations on equal footing become feasible. We note, however, that we have set out to provide such a program package.[4]

7. CONCLUSIONS

In this work, we provided a general description for automated atomistic reaction network exploration algorithms. Owing to the huge diversity and heterogeneity of the tasks, we considered it necessary to define core concepts, targets, and challenges in order to make this fast-growing field accessible to assessment and validation and to identify weak spots in existing schemes. Although a detailed rating of existing schemes has been beyond the scope of this work for various reasons, of which the limited accessibility of implementations is an important one, the present overview may still be well suited as a general guideline toward the application range and capabilities of existing algorithms.

We defined three basic exploration patterns for chemical reaction networks that depend on the scientific purpose for which they are designed. In connection with a basic nomenclature presented afterward, we then introduced two fidelity measures related to the breadth and depth of a network. These are graph accuracy, i.e., the correct identification and connectivity of nodes, and node accuracy, i.e., the correct description of the set of structures that represent a compound. We continued to classify prototypical tasks for network exploration algorithms. These include a high level of automation, error diagnostics and automated uncertainty handling, network data transferability, and eventually kinetic modeling. Finally, we proposed a transferable comparative scheme

---

[4]The `scine` developers provide software for chemical interaction networks (available at **https://scine.ethz.ch/**).

of reaction space exploration software based on the concepts and targets elaborated on in this work. Key measures will be node and graph accuracy; software extensibility, automation, immersion, and user-friendliness; data transferability; and of course, the computational cost associated with a given algorithm/software.

We are certain that such a multidimensional ranking will become decisive in future work on reaction network exploration algorithms as it can advance the field by clear categorization of new exploration protocols in the context of existing ones. Detailed quantitative measures can then be provided for existing networks, and new schemes can emerge, once balanced benchmark networks become available, that represent the variety of algorithmic features highlighted in this work.

SUMMARY POINTS

1. Predictive theoretical work on molecular reactivity and function will require in all but the simplest cases the study of a huge amount of molecular structures and their relationships.

2. Automated procedures are mandatory for this task and recent developments point to fruitful directions for such approaches.

3. Reaction-space exploration approaches must address a wide array of requirements ranging from stable exploration procedures for advanced atomistic modeling to uncertainty quantification for error assessment and cure, to new visualization and immersion software and hardware to interact with complex networks of many thousand nodes.

4. The comparison of exploration algorithms and software is a multidimensional task that requires careful assessment of the pros and cons regarding their theoretical background, efficient and stable implementation, and overall software engineering.

FUTURE ISSUES

1. Future developments of algorithms for chemical (reaction) space exploration will require a very broad theoretical basis in order to cope with the many different scenarios that one may encounter in real-world molecular processes.

2. Such algorithmic developments will eventually demand a high level of sophistication of software engineering and integration in order to move on from advanced tools for specific problems to general tool boxes for explorations of molecular processes at the nanoscale.

3. Eventually, fully automated computational chemistry software will become a peer-to-data-driven operation with respect to molecular science in research and education if it can act autonomously on arguments and questions raised in natural language by the operator.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

LITERATURE CITED

1. Corey EJ, Cheng XM. 1995. *The Logic of Chemical Synthesis*. New York: Wiley. 1st ed.
2. Broadbelt LJ, Stark SM, Klein MT. 1994. Computer generated pyrolysis modeling: on-the-fly generation of species, reactions, and rates. *Ind. Eng. Chem. Res.* 33:790–99
3. Broadbelt LJ, Stark S, Klein M. 1996. Computer generated reaction modelling: decomposition and encoding algorithms for determining species uniqueness. *Comput. Chem. Eng.* 20:113–29
4. Broadbelt LJ, Pfaendtner J. 2005. Lexicography of kinetic modeling of complex reaction networks. *AIChE J.* 51:2112–21
5. Schwaller P, Laino T, Gaudin T, Bolgar P, Hunter CA, et al. 2019. Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction. *ACS Cent. Sci.* 5:1572–83
6. Schwaller P, Gaudin T, Lnyi D, Bekas C, Laino T. 2018. "Found in translation": predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models. *Chem. Sci.* 9:6091–98
7. Segler MHS, Preuss M, Waller MP. 2018. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* 555:604–10
8. Pensak DA, Corey EJ. 1977. LHASA—logic and heuristic applied to synthetic analysis. In *Computer-Assisted Organic Synthesis*, Vol. 61, ed. WT Wipke, WJ Howe, pp. 1–32. Washington, DC: Am. Chem. Soc.
9. Ihlenfeldt WD, Gasteiger J. 1996. Computer-assisted planning of organic syntheses: the second generation of programs. *Angew. Chem. Int. Ed.* 34:2613–33
10. Gothard CM, Soh S, Gothard NA, Kowalczyk B, Wei Y, et al. 2012. Rewiring chemistry: algorithmic discovery and experimental validation of one-pot reactions in the network of organic chemistry. *Angew. Chem. Int. Ed.* 51:7922–27
11. Kowalik M, Gothard CM, Drews AM, Gothard NA, Weckiewicz A, et al. 2012. Parallel optimization of synthetic pathways within the network of organic chemistry. *Angew. Chem. Int. Ed.* 51:7928–32
12. Segler MHS, Waller MP. 2017. Neural-symbolic machine learning for retrosynthesis and reaction prediction. *Chem. Eur. J.* 23:5966–71
13. Coley CW, Rogers L, Green WH, Jensen KF. 2017. Computer-assisted retrosynthesis based on molecular similarity. *ACS Cent. Sci.* 3:1237–45
14. Coley CW, Barzilay R, Jaakkola TS, Green WH, Jensen KF. 2017. Prediction of organic reaction outcomes using machine learning. *ACS Cent. Sci.* 3:434–43
15. Segler MHS, Waller MP. 2017. Modelling chemical reasoning to predict and invent reactions. *Chem. Eur. J.* 23:6118–28
16. Coley CW, Green WH, Jensen KF. 2018. Machine learning in computer-aided synthesis planning. *Acc. Chem. Res.* 51:1281–89
17. Simm GN, Vaucher AC, Reiher M. 2019. Exploration of reaction pathways and chemical transformation networks. *J. Phys. Chem. A* 123:385–99
18. Sameera WMC, Maeda S, Morokuma K. 2016. Computational catalysis using the artificial force induced reaction method. *Acc. Chem. Res.* 49:763–73
19. Dewyer AL, Argüelles AJ, Zimmerman PM. 2018. Methods for exploring reaction space in molecular systems. *WIREs Comput. Mol. Sci.* 8:e1354
20. Habershon S. 2016. Automated prediction of catalytic mechanism and rate law using graph-based reaction path sampling. *J. Chem. Theory Comput.* 12:1786–98
21. Kim Y, Kim JW, Kim Z, Kim WY. 2018. Efficient prediction of reaction paths through molecular graph and reaction network analysis. *Chem. Sci.* 9:825–35
22. Ismail I, Stuttaford-Fowler HBVA, Ochan Ashok C, Robertson C, Habershon S. 2019. Automatic proposal of multistep reaction mechanisms using a graph-driven search. *J. Phys. Chem. A* 123:3407–17

23. Bergeler M, Simm GN, Proppe J, Reiher M. 2015. Heuristics-guided exploration of reaction mechanisms. *J. Chem. Theory Comput.* 11:5712–22

24. Simm GN, Reiher M. 2017. Context-driven exploration of complex chemical reaction networks. *J. Chem. Theory Comput.* 13:6108–19

25. Grimmel S, Reiher M. 2019. The electrostatic potential as a descriptor for the protonation propensity in automated exploration of reaction mechanisms. *Faraday Discuss.* 220:443–63

26. Rappoport D, Galvin CJ, Zubarev DY, Aspuru-Guzik A. 2014. Complex chemical reaction networks from heuristics-aided quantum chemistry. *J. Chem. Theory Comput.* 10:897–907

27. Rappoport D, Aspuru-Guzik A. 2019. Predicting feasible organic reaction pathways using heuristically aided quantum chemistry. *J. Chem. Theory Comput.* 15(7):4099–112

28. Kim Y, Choi S, Kim WY. 2014. Efficient basin-hopping sampling of reaction intermediates through molecular fragmentation and graph theory. *J. Chem. Theory Comput.* 10:2419–26

29. Habershon S. 2015. Sampling reactive pathways with random walks in chemical space: applications to molecular dissociation and catalysis. *J. Chem. Phys.* 143:094106

30. Ohno K, Maeda S. 2004. A scaled hypersphere search method for the topography of reaction pathways on the potential energy surface. *Chem. Phys. Lett.* 384:277–82

31. Maeda S, Morokuma K. 2010. Communications: a systematic method for locating transition structures of $A + B \rightarrow X$ type reactions. *J. Chem. Phys.* 132:241102

32. Maeda S, Ohno K, Morokuma K. 2013. 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.* 15:3683–701

33. Maeda S, Harabuchi Y, Takagi M, Saita K, Suzuki K, et al. 2018. Implementation and performance of the artificial force induced reaction method in the GRRM17 program. *J. Comput. Chem.* 39:233–51

34. Zimmerman PM. 2015. Single-ended transition state finding with the growing string method. *J. Comput. Chem.* 36:601–11

35. Dewyer AL, Zimmerman PM. 2017. Finding reaction mechanisms, intuitive or otherwise. *Org. Biomol. Chem.* 15:501–4

36. Wang LP, Titov A, McGibbon R, Liu F, Pande VS, Martínez TJ. 2014. Discovering chemistry with an ab initio nanoreactor. *Nat. Chem.* 6:1044

37. Yang M, Zou J, Wang G, Li S. 2017. Automatic reaction pathway search via combined molecular dynamics and coordinate driving method. *J. Phys. Chem. A* 121:1351–61

38. Huber T, Torda AE, van Gunsteren WF. 1994. Local elevation: a method for improving the searching properties of molecular dynamics simulation. *J. Comput.-Aided Mol. Des.* 8:695–708

39. Laio A, Parrinello M. 2002. Escaping free-energy minima. *PNAS* 99:12562–66

40. Grimme S. 2019. Exploration of chemical compound, conformer, and reaction space with meta-dynamics simulations based on tight-binding quantum chemical calculations. *J. Chem. Theory Comput.* 15:2847–62

41. Rizzi V, Mendels D, Sicilia E, Parrinello M. 2019. Blind search for complex chemical pathways using harmonic linear discriminant analysis. *J. Chem. Theory Comput.* 15(8):4507–15

42. Martínez-Núñez E. 2015. An automated method to find transition states using chemical dynamics simulations. *J. Comput. Chem.* 36:222–34

43. Varela JA, Vázquez SA, Martínez-Núñez E. 2017. An automated method to find reaction mechanisms and solve the kinetics in organometallic catalysis. *Chem. Sci.* 8:3843–51

44. Yang M, Yang L, Wang G, Zhou Y, Xie D, Li S. 2018. Combined molecular dynamics and coordinate driving method for automatic reaction pathway search of reactions in solution. *J. Chem. Theory Comput.* 14:5787–96

45. Debnath J, Invernizzi M, Parrinello M. 2019. Enhanced sampling of transition states. *J. Chem. Theory Comput.* 15:2454–59

46. Grambow CA, Jamal A, Li YP, Green WH, Zádor J, Suleimanov YV. 2018. Unimolecular reaction pathways of a γ-ketohydroperoxide from combined application of automated reaction discovery methods. *J. Am. Chem. Soc.* 140:1035–48

47. Maeda S, Harabuchi Y. 2019. On benchmarking of automated methods for performing exhaustive reaction path search. *J. Chem. Theory Comput.* 15:2111–15

48. Green WH, Moore CB, Polik WF. 1992. Transition states and rate constants for unimolecular reactions. *Annu. Rev. Phys. Chem.* 43:591–626

49. Susnow RG, Dean AM, Green WH, Peczak P, Broadbelt LJ. 1997. Rate-based construction of kinetic models for complex systems. *J. Phys. Chem. A* 101:3731–40

50. Sumathi R, Green WH Jr. 2002. A priori rate constants for kinetic modeling. *Theor. Chem. Acc.* 108:187–213

51. Prinz JH, Wu H, Sarich M, Keller B, Senne M, et al. 2011. Markov models of molecular kinetics: generation and validation. *J. Chem. Phys.* 134:174105

52. Sabbe MK, Reyniers MF, Reuter K. 2012. First-principles kinetic modeling in heterogeneous catalysis: an industrial perspective on best-practice, gaps and needs. *Catal. Sci. Technol.* 2:2010–24

53. Bowman GR, Pande VS, Noé F, eds. 2014. *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*. Dordrecht, Neth: Springer Sci. Bus. Med.

54. Van de Vijver R, Vandewiele NM, Bhoorasingh PL, Slakman BL, Seyedzadeh Khanshan F, et al. 2015. 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.* 47:199–231

55. Gao CW, Allen JW, Green WH, West RH. 2016. Reaction mechanism generator: automatic construction of chemical kinetic mechanisms. *Comput. Phys. Commun.* 203:212–25

56. Wu H, Paul F, Wehmeyer C, Noé F. 2016. Multiensemble Markov models of molecular thermodynamics and kinetics. *PNAS* 113:E3221–30

57. Proppe J, Husch T, Simm GN, Reiher M. 2016. Uncertainty quantification for quantum chemical models of complex reaction networks. *Faraday Discuss.* 195:497–520

58. Döpking S, Matera S. 2017. Error propagation in first-principles kinetic Monte Carlo simulation. *Chem. Phys. Lett.* 674:28–32

59. Han K, Green WH. 2018. A fragment-based mechanistic kinetic modeling framework for complex systems. *Ind. Eng. Chem. Res.* 57:14022–30

60. Vereecken L, Aumont B, Barnes I, Bozzelli J, Goldman M, et al. 2018. Perspective on mechanism development and structure-activity relationships for gas-phase atmospheric chemistry. *Int. J. Chem. Kinet.* 50:435–69

61. Scherer MK, Husic BE, Hoffmann M, Paul F, Wu H, No F. 2019. Variational selection of features for molecular kinetics. *J. Chem. Phys.* 150:194108

62. Andersen M, Panosetti C, Reuter K. 2019. A practical guide to surface kinetic Monte Carlo simulations. *Front. Chem.* 7:202

63. Proppe J, Reiher M. 2019. Mechanism deduction from noisy chemical reaction networks. *J. Chem. Theory Comput.* 15:357–70

64. Cavallotti C, Pelucchi M, Georgievskii Y, Klippenstein SJ. 2019. EStokTP: electronic structure to temperature- and pressure-dependent rate constants-A code for automatically predicting the thermal kinetics of reactions. *J. Chem. Theory Comput.* 15:1122–45

65. Eyring H. 1935. The activated complex in chemical reactions. *J. Chem. Phys.* 3:107–15

66. Kramers H. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7:284–304

67. Hänggi P, Talkner P, Borkovec M. 1990. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* 62:251–341

68. Weymuth T, Reiher M. 2014. Inverse quantum chemistry: concepts and strategies for rational compound design. *Int. J. Quantum Chem.* 114:823–37

69. Sanchez-Lengeling B, Aspuru-Guzik A. 2018. Inverse molecular design using machine learning: generative models for matter engineering. *Science* 361:360–65

70. Freeze JG, Kelly HR, Batista VS. 2019. Search for catalysts by inverse design: artificial intelligence, mountain climbers, and alchemists. *Chem. Rev.* 119:6595–612

71. Collins KD, Rühling A, Glorius F. 2014. Application of a robustness screen for the evaluation of synthetic organic methodology. *Nat. Prot.* 9:1348

72. Mayer I. 1983. Charge, bond order and valence in the ab initio SCF theory. *Chem. Phys. Lett.* 97:270–74

73. Karpen ME, Tobias DJ, Brooks CL. 1993. Statistical clustering techniques for the analysis of long molecular dynamics trajectories: analysis of 2.2-ns trajectories of YPGDV. *Biochemistry* 32:412–20

74. Shenkin PS, McDonald DQ. 1994. Cluster analysis of molecular conformations. *J. Comput. Chem.* 15:899–916

75. Jain AK, Murty MN, Flynn PJ. 1999. Data clustering: a review. *ACM Comput. Surv.* 31:264–323

76. Shao J, Tanner SW, Thompson N, Cheatham TE. 2007. Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.* 3:2312–34

77. Keller B, Daura X, van Gunsteren WF. 2010. Comparing geometric and kinetic cluster algorithms for molecular simulation data. *J. Chem. Phys.* 132:074110

78. Singhal N, Snow CD, Pande VS. 2004. Using path sampling to build better Markovian state models: predicting the folding rate and mechanism of a tryptophan zipper beta hairpin. *J. Chem. Phys.* 121:415–25

79. Bowman GR, Huang X, Pande VS. 2009. Using generalized ensemble simulations and Markov state models to identify conformational states. *Methods* 49:197–201

80. Li Y, Dong Z. 2016. Effect of clustering algorithm on establishing Markov state model for molecular dynamics simulations. *J. Chem. Inf. Model.* 56:1205–15

81. Vaucher AC, Reiher M. 2016. Molecular propensity as a driver for explorative reactivity studies. *J. Chem. Inf. Model.* 56:1470–78

82. Simm GN, Proppe J, Reiher M. 2017. Error assessment of computational models in chemistry. *CHIMIA* 71:202–8

83. Simm GN, Reiher M. 2018. Error-controlled exploration of chemical reaction networks with Gaussian processes. *J. Chem. Theory Comput.* 14:5238–48

84. Wang LP, Song C. 2019. Car-Parrinello monitor for more robust Born-Oppenheimer molecular dynamics. *J. Chem. Theory Comput.* 15(8):4454–67

85. Vaucher AC, Reiher M. 2017. Steering orbital optimization out of local minima and saddle points toward lower energy. *J. Chem. Theory Comput.* 13:1219–28

86. Hawkins PCD. 2017. Conformation generation: the state of the art. *J. Chem. Inf. Model.* 57:1747–56

87. Klamt A. 1995. Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem.* 99:2224–35

88. Grambow CA, Li YP, Green WH. 2019. Accurate thermochemistry with small data sets: a bond additivity correction and transfer learning approach. *J. Phys. Chem. A* 123(27):5826–35

89. Sidler D, Schwaninger A, Riniker S. 2016. Replica exchange enveloping distribution sampling (re-eds): a robust method to estimate multiple free-energy differences from a single simulation. *J. Chem. Phys.* 145:154114

90. Sidler D, Cristfol-Clough M, Riniker S. 2017. Efficient round-trip time optimization for replica-exchange enveloping distribution sampling (re-eds). *J. Chem. Theory Comput.* 13:3020–30

91. Gomes ASP, Jacob CR. 2012. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* 108:222–77

92. Jacob CR, Neugebauer J. 2014. Subsystem density-functional theory. *WIREs Comput. Mol. Sci.* 4:325–62

93. Wesolowski TA, Shedge S, Zhou X. 2015. Frozen-density embedding strategy for multilevel simulations of electronic structure. *Chem. Rev.* 115:5891–928

94. Lee SJR, Welborn M, Manby FR, Miller TF. 2019. Projection-based wavefunction-in-DFT embedding. *Acc. Chem. Res.* 52:1359–68

95. Husch T, Vaucher AC, Reiher M. 2018. Semiempirical molecular orbital models based on the neglect of diatomic differential overlap approximation. *Int. J. Quantum Chem.* 118:e25799

96. Parr RG, Yang W. 1994. *Density-Functional Theory of Atoms and Molecules*. New York: Oxford Univ. Press

97. Weymuth T, Couzijn EPA, Chen P, Reiher M. 2014. New benchmark set of transition-metal coordination reactions for the assessment of density functionals. *J. Chem. Theory Comput.* 10:3092–103

98. Simm GN, Reiher M. 2016. Systematic error estimation for chemical reaction energies. *J. Chem. Theory Comput.* 12:2762–73

99. Husch T, Freitag L, Reiher M. 2018. Calculation of ligand dissociation energies in large transition-metal complexes. *J. Chem. Theory Comput.* 14:2456–68

100. Pernot P, Civalleri B, Presti D, Savin A. 2015. Prediction uncertainty of density functional approximations for properties of crystals with cubic symmetry. *J. Phys. Chem. A* 119:5288–304

101. Pernot P. 2017. The parameter uncertainty inflation fallacy. *J. Chem. Phys.* 147:104102

102. Proppe J, Reiher M. 2017. Reliable estimation of prediction uncertainty for physicochemical property models. *J. Chem. Theory Comput.* 13:3297–317

103. Ma Q, Werner HJ. 2018. Explicitly correlated local coupled-cluster methods using pair natural orbitals. *WIREs Comput. Mol. Sci.* 8:e1371

104. Stein CJ, Reiher M. 2016. Automated selection of active orbital spaces. *J. Chem. Theory Comput.* 12:1760–71

105. Stein CJ, von Burg V, Reiher M. 2016. The delicate balance of static and dynamic electron correlation. *J. Chem. Theory Comput.* 12:3764–73

106. Stein CJ, Reiher M. 2019. autoCAS: a program for fully automated multiconfigurational calculations. *J. Comput. Chem.* 40(25):2216–26

107. Manby FR, Stella M, Goodpaster JD, Miller TF. 2012. A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* 8:2564–68

108. Tamukong PK, Khait YG, Hoffmann MR. 2014. Density differences in embedding theory with external orbital orthogonality. *J. Phys. Chem. A* 118:9182–200

109. Hégely B, Nagy PR, Ferenczy GG, Kállay M. 2016. Exact density functional and wave function embedding schemes based on orbital localization. *J. Chem. Phys.* 145:064107

110. Mühlbach AH, Reiher M. 2018. Quantum system partitioning at the single-particle level. *J. Chem. Phys.* 149:184104

111. Proppe J, Gugler S, Reiher M. 2019. Gaussian process-based refinement of dispersion corrections. *J. Chem. Theory Comput.* 15(11):6046–60

112. Haag MP, Vaucher AC, Bosson M, Redon S, Reiher M. 2014. Interactive chemical reactivity exploration. *ChemPhysChem* 15:3301–19

113. Haag MP, Reiher M. 2014. Studying chemical reactivity in a virtual environment. *Faraday Discuss.* 169:89–118

114. Luehr N, Jin AGB, Martínez TJ. 2015. Ab initio interactive molecular dynamics on graphical processing units (GPUs). *J. Chem. Theory Comput.* 11(10):4536–44

115. O'Connor M, Deeks HM, Dawn E, Metatla O, Roudaut A, et al. 2018. Sampling molecular conformations and dynamics in a multiuser virtual reality framework. *Sci. Adv.* 4:eaat2731

116. Amabilino S, Bratholm LA, Bennie SJ, Vaucher AC, Reiher M, Glowacki DR. 2019. Training neural nets to learn reactive potential energy surfaces using interactive quantum chemistry in virtual reality. *J. Phys. Chem. A* 123:4486–99

117. Lu Y. 2017. The "OK, Molly" chemistry. *Acc. Chem. Res.* 50:647–51

118. Aspuru-Guzik A, Lindh R, Reiher M. 2018. The matter simulation (r)evolution. *ACS Cent. Sci.* 4:144–52

119. Fast E, Chen B, Mendelsohn J, Bassen J, Bernstein M. 2018. Iris: a conversational agent for complex tasks. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. New York: ACM

120. Lejaeghere K, Bihlmayer G, Björkman T, Blaha P, Blügel S, et al. 2016. Reproducibility in density functional theory calculations of solids. *Science* 351:aad3000

121. Glowacki DR, Liang CH, Morley C, Pilling MJ, Robertson SH. 2012. MESMER: an open-source master equation solver for multi-energy well reactions. *J. Phys. Chem. A* 116:9545–60

122. Kee R, Miller J, Jefferson T. 1980. *CHEMKIN: a general-purpose, problem-independent, transportable, FORTRAN chemical kinetics code package*. Tech. Rep. SAND80-8003, Sandia Labs

123. Miller WH. 1993. Beyond transition-state theory: a rigorous quantum theory of chemical reaction rates. *Acc. Chem. Res.* 26:174–81

124. Truhlar DG, Garrett BC, Klippenstein SJ. 1996. Current status of transition-state theory. *J. Phys. Chem.* 100:12771–800

125. Pollak E, Talkner P. 2005. Reaction rate theory: What it was, where is it today, and where is it going? *Chaos* 15:026116

126. Garrett BC, Truhlar DG. 2005. Variational transition state theory. In *Theory and Applications of Computational Chemistry*, ed. CE Dykstra, G Frenking, KS Kim, GE Scuseria, pp. 67–87. Amsterdam: Elsevier
127. R. Soc. Chem. 2017. *Discussion Volume from the Faraday Discussion on Reaction Rate Theory, September 19–21, 2016, Cambridge, UK*. Cambridge, UK: R. Soc. Chem.
128. Richardson JO. 2018. Perspective: ring-polymer instanton theory. *J. Chem. Phys.* 148:200901
129. Reilly AM, Cooper RI, Adjiman CS, Bhattacharya S, Boese AD, et al. 2016. Report on the sixth blind test of organic crystal structure prediction methods. *Act. Cryst. B* 72:439–59
130. Rizzi A, Murkli S, McNeill JN, Yao W, Sullivan M, et al. 2018. Overview of the SAMPL6 host–guest binding affinity prediction challenge. *J. Comput.-Aided Mol. Des.* 32:937–63
131. Maeda S, Morokuma K. 2012. Toward predicting full catalytic cycle using automatic reaction path search method: a case study on HCo(CO)$_3$-catalyzed hydroformylation. *J. Chem. Theory Comput.* 8:380–85