

Manuscript version: Published Version

The version presented in WRAP is the published version (Version of Record).

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/166663>

How to cite:

The repository item page linked to above, will contain details on accessing citation guidance from the publisher.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Successes and challenges in using machine-learned activation energies in kinetic simulations

Cite as: J. Chem. Phys. 157, 014109 (2022); doi: 10.1063/5.0096027

Submitted: 14 April 2022 • Accepted: 13 June 2022 •

Published Online: 5 July 2022



View Online



Export Citation



CrossMark

I. Ismail,¹ C. Robertson,²  and S. Habershon^{1,a)} 

AFFILIATIONS

¹ Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom

² School of Chemistry, Heriot-Watt University, Edinburgh, EH14 4AS United Kingdom

Note: This paper is part of the JCP Special Topic on Chemical Design by Artificial Intelligence.

^{a)} Author to whom correspondence should be addressed: S.Habershon@warwick.ac.uk

ABSTRACT

The prediction of the thermodynamic and kinetic properties of chemical reactions is increasingly being addressed by machine-learning (ML) methods, such as artificial neural networks (ANNs). While a number of recent studies have reported success in predicting chemical reaction activation energies, less attention has been focused on how the accuracy of ML predictions filters through to predictions of macroscopic observables. Here, we consider the impact of the uncertainty associated with ML prediction of activation energies on observable properties of chemical reaction networks, as given by microkinetics simulations based on ML-predicted reaction rates. After training an ANN to predict activation energies, given standard molecular descriptors for reactants and products alone, we performed microkinetics simulations of three different prototypical reaction networks: formamide decomposition, aldol reactions, and decomposition of 3-hydroperoxypropanal. We find that the kinetic modeling predictions can be in excellent agreement with corresponding simulations performed with *ab initio* calculations, but this is dependent on the inherent energetic landscape of the networks. We use these simulations to suggest some guidelines for when ML-based activation energies can be reliable and when one should take more care in applications to kinetics modeling.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0096027>

I. INTRODUCTION

Chemical reaction networks (CRNs) are a powerful tool to understand how *macroscopic*, experimentally measurable properties, such as reaction rates, reaction orders, product selectivities, and branching ratios, emerge from *microscopic* characteristics, such as reaction activation energies and relative molecular energetics.^{1–19} CRNs comprise a set of nodes (or vertices) representing discrete molecular species (or related collections thereof) and a set of edges connecting nodes that represent the chemical reactions interconnecting different chemical species. The entire CRN is typically characterized by defining the connectivity, the relative energies of the different nodes, and the reaction rates of each chemical reaction; most commonly, one defines the activation energies of each reaction and assumes that transition-state theory (TST) or related analytical methods can be used to determine reaction rates.^{17,20–22} Once the CRN characteristics are fully defined, microkinetics simulations can be used to predict the transient behavior for a given

set of initial species concentrations, allowing access to simulations that can explain observed overall rate laws or product distributions for complex chemical systems.^{14,18,23–28}

CRNs can be created in different ways, for example, using experimentally measured rate constants, activation energies from *ab initio* electronic structure calculations, or a mixture of both methods; in this article, we focus exclusively on computational methods for generating CRNs. In such cases, the procedure for generating CRNs can be summarized as (i) generating species lists and characterizing their relative energies, (ii) generating reaction lists defining the set of allowed chemical reactions, (iii) calculating activation energies for all chemical reactions, and (iv) calculating reaction rates for all reactions.

The last few decades have seen enormous progress being made in steps (i) and (ii), namely, automated computational generation of species and chemical reactions. For example, the reaction mechanism generator (RMG^{1,29–32}) has been used fruitfully to study a

wide range of complex CRNs, such as unimolecular decomposition and pyrolysis of organic species. The AutoMeKin^{33–37} code employs molecular dynamics (MD) simulations to drive the search for chemical reactions and transition states (TSs) and has similarly been employed to study reactions ranging from organometallic catalysis to ozonolysis. A number of other reaction generation schemes have been proposed,^{2,3,7,38–43} such as graph-based sampling methods that drive chemical reactions in the discretized space formed from molecular adjacency matrices,^{10,11} MD-type schemes that treat reaction paths as dynamic objects with constrained end-points,^{12,44} and *ab initio* MD schemes that employ artificial pistons to periodically force reactive chemical species into close enough proximity to enable chemical reactions to occur.⁴⁵ A number of recent articles have provided excellent overviews of the emerging available methods for automated generation of reactive species and reactions for CRN construction, and this continues to be a continued growth area in computational chemistry.^{7,40}

In accurate modeling of CRNs, we are then left with the challenge of accurate evaluation of activation energies and reaction rates [steps (iii) and (iv)] as the key challenges. As noted above, TST and related analytical methods are almost universally employed in calculating reactions rates for computational investigations of CRNs; this has the advantage of being computationally straightforward, typically requiring information about the relative energies of the TS and the reactive minima. Furthermore, the underlying assumptions of TST, such as the “no recrossing” assumption, typically introduce relatively small errors when compared to errors in calculated activation energies; this is a simple result of the fact that TST rates are exponentially dependent on calculated activation energies. As a result, the accurate determination of the TS and calculation of the corresponding activation energy remain an important challenge in computational CRN analysis.^{7,33,36,37,39,46,47} Furthermore, we note that many CRNs, notably those related to combustion processes or atmospheric chemistry in urban environments, can contain enormous numbers of molecular reactive species and chemical reactions;^{30,48–53} in such cases, the sheer size of the computational task of evaluating activation energies is a bottle-neck, particularly if one requires accurate *ab initio* evaluations.

The challenges of accurate prediction of activation barriers for large datasets of chemical reactions in an automated, high-throughput manner are an ideal setting for machine-learning (ML). Similar to methods for computational reaction discovery, the last decade or so has seen a growth in strategies for ML prediction of activation energies for chemical reactions.^{54–58} In large part, this activity has been driven by the increasing availability of high-quality and publicly available datasets containing large numbers (typically several thousand or more) of individual chemical reactions described through reactant and product structures and the corresponding activation energies and reaction energetics.^{54,59,60}

Given such a training set, each reaction is typically characterized (or vectorized) using one of a wide range of possible descriptors aimed at capturing the chemically relevant features of the reaction, such as connectivity and atom-types of reactive sites, type of reaction, and thermodynamic properties such as reaction energy changes [following widely known relations such as the Bronsted–Evans–Polanyi (BEP) relation connecting activation energy and reaction energetics^{58,61–65}]. In this context, a wide range of descriptors have been employed, including sub-fragment-based

strategies, such as Morgan fingerprints, connectivity-based schemes, or mixtures of descriptors containing both connectivity and geometric information.^{59,66–70} Most commonly, in order to avoid the computational expense and challenge of locating TSs for all relevant reactions in a CRN, reaction descriptors are calculated using information for just the reactant and product molecular species alone.

After descriptor calculation for the target dataset, ML strategies, such as artificial neural networks (ANNs) or kernel regression methods,⁷¹ can be used to generate predictive models to approximate activation energies, given input reactant and product structures in the form of simplified molecular-input line-entry system (SMILES) strings or molecular geometries. Subsequently, the trained ANN or regression method can be used to predict activation barriers for related chemical reactions; as such, it is clear that ML strategies, given appropriate training data, can be used to dramatically accelerate the calculation of activation energies in CRNs, circumventing the necessity of computationally demanding *ab initio* calculations and TS-location schemes.

The general strategy outlined above has been successfully employed in studies of a variety of different chemical systems, ranging from organic chemical reactions to heterogeneous catalysis. From these previous studies, performed with a variety of different descriptors, regression strategies, and training datasets, it is found that the typical accuracy currently attainable through ML strategies for activation energy prediction is in the range 2–6 kcal mol⁻¹ [relative to the “correct” barriers typically given by density functional theory (DFT)].

Following our previous work on reaction discovery^{10–12,44} and our ambition to integrate our reaction discovery schemes with automated generation and microkinetics modeling of CRNs, the purpose of this article is to test the extent to which current ML-predicted activation energies are sufficiently accurate to reproduce emergent CRN kinetics. Here, we perform microkinetics simulations of three different examples of CRNs, chosen to represent different characteristics; in each case, we compare the kinetics for CRNs generated with (i) *ab initio*-predicted activation energies and (ii) ML-predicted activation energies in order to assess the conditions and CRN characteristics under which ML predictions of activation energies are “good enough.” By comparing CRN microkinetics for both ML activation energy prediction and *ab initio* barriers, we hope that these simulations might help further establish the utility of ML in challenging simulations of CRNs.

In Sec. II, we first describe our approach to ANN fitting of activation energies using Morgan fingerprints and ANNs. In Sec. III, we subsequently demonstrate the characteristics and accuracy of the trained ANN before deploying the ANN activation energy predictions in microkinetics simulations of formamide decomposition,⁷² aldol reactions,⁴³ and decomposition of 3-hydroperoxypropanal.³² Finally, we conclude by highlighting what our results suggest for the further deployment of ML in CRNs.

II. THEORY

In this section, we begin by outlining our approach to activation energy prediction using ANNs; as noted above, our regression strategy here is comparable to a number of previous studies and employs readily available molecular descriptors and training

datasets. In other words, by definition, our approach is representative of the broad state-of-play in using ANNs to predict activation energies for typical organic reactions.

After outlining our ANN strategy, we briefly describe how the predicted activation energies feed into microkinetics simulations of CRNs; we also highlight how predicted uncertainties are incorporated into our microkinetics predictions.

A. ANN prediction of activation energies

We begin by describing the procedure followed here to train an ANN to predict activation energies of organic chemical reactions. In general, we follow similar strategies as have been deployed previously, focusing on using Morgan (extended-connectivity) fingerprints^{66,67} as molecular descriptors for reactants and products and using a well-known organic chemistry DFT dataset for training and testing.⁶⁰ As noted above, the point of this article is to investigate whether such “standard” approaches are good enough for predictive microkinetics modeling and what can be learned from this approach.

In this article, we follow a path that has been broadly employed previously in seeking to predict activation energies for chemical reactions with defined reactants and products (Fig. 1); importantly, the ANN will be trained using only reaction end-point data, including connectivity properties for the reaction end-points and the *ab initio*-calculated reaction energy change. Following previous work, these properties are encoded in a feature vector (described in the following); subsequently, a feature vector for the reaction is obtained as the difference between the product and reactant feature vectors, providing an encoding of the structural and energetic changes induced by the reaction itself. This final reaction feature vector is then used as input to an ANN with the aim of predicting the corresponding activation energy. A schematic overview of the general ANN setup is shown in Fig. 1.

We note here that a number of previous investigations have similarly deployed ANNs or related regression strategies to predict activation energies for similar chemical reactions.^{54–56,73,74} In a recent article, Singh and co-workers⁷³ developed a multi-feature neural network model and discovered that they were able to predict activation barriers with mean absolute errors (MAEs) as low as

5.07 kcal mol⁻¹. Similarly, Choi and co-workers⁵⁵ exploited molecular fingerprints and decision trees to determine activation energies, achieving a mean absolute error of 2.0 kcal mol⁻¹. A survey of these, and other, previous investigations reveals that it is broadly possible to obtain ANNs that can predict activation energies for diverse chemical reactions with a typical root-mean-square error (RMSE) error of 2.8 kcal mol⁻¹ and mean absolute error (MAE) of 2.5 kcal mol⁻¹. A key aim of this article is to generate an ANN with comparable predictive capability and then to explore how useful this level of accuracy is in predicting emergent kinetics of CRNs. In passing, we note that we have also tried to employ Gaussian process regression (GPR) to fit activation energies using the same dataset and fingerprints as described in the following; however, we found that the predictive performance was typically lower than that offered by ANNs for this particular application, and we chose to focus here exclusively on ANNs.

In our ANN approach, we use standard ANN architectures, datasets, and feature vectors, which are widely available; these different aspects are described in Secs. II A 1–II A 3.

1. Dataset for ANN training

As described above, our main simulation focus in the following is in modeling three different CRNs with different overall structures; however, each of these CRNs is based around organic chemical reactions with no transition metals present. In order to train an ANN to predict activation energies for these systems, we require a training dataset that spans a similar range of chemical reaction space. Recently, Grambow and co-workers⁶⁰ have published such a dataset based solely on DFT calculations. This dataset comprises over 16 000 gas-phase organic molecular reactions for species containing up to seven heteroatoms (C, N, and O); for each reaction, the dataset contains the reaction energy change and the TS energy (as determined by the growing-string method), all calculated at the DFT B97-D3/def2-mSVP level. We note that a comparable dataset based on the alternative DFT function ω B97X-D3 is also available for organic chemical reactions; however, the B97-D3 dataset employed here is larger and was therefore viewed as more appropriate for our purpose. Furthermore, we note that the fitting and testing methodology discussed here is not particularly tied to any specific functional.

To make full use of this dataset, we incorporate both the forward and backward mechanisms, resulting in a database of over 32 000 organic reactions, which can be used for ANN training and testing. However, we emphasize that while our training set includes a mixture of forward and backward reactions, the test set only include examples of forward reactions; this approach prevents the same reactions being included in both training- and test-sets, enabling more reliable assessment of ANN accuracy.

2. Molecular descriptors

To effectively learn activation energies, chemical reaction features must be encoded in a form suitable for an ANN; we need a compact descriptor set that captures the characteristics of the reactants, products, and associated energy changes for a given chemical reaction. While recent work has developed graph-convolutional neural networks (GCNNs)⁷⁵ and quantum-chemically derived reactivity descriptors, the approach we take here is to focus on simpler, readily available molecular descriptors that can be directly calculated

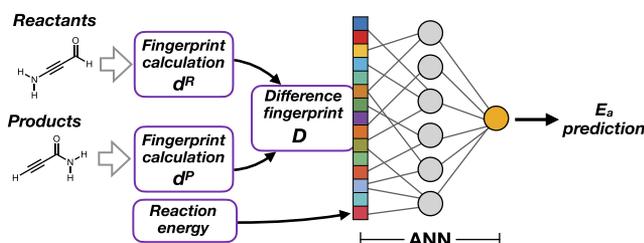


FIG. 1. Schematic representation of ML prediction for activation energies used in this article. For a given set of reactants and products, separate molecular fingerprints are calculated, d^R and d^P , and then combined into a reaction-difference fingerprint D [Eq. (1)]. With the addition of the reaction energy, this difference fingerprint is then used as the input vector to an ANN with the output giving an estimate of the activation energy. The weights in the ANN are trained using a database of entries containing reactant/product SMILES strings and corresponding DFT activation energies.

from input reactant and product structures without further external input.

Connectivity-based models offer a simple, yet robust representation of chemical structures; as noted above, these can be subtracted to generate a “reaction fingerprint.” The descriptor we use here is the Morgan (or extended-connectivity) fingerprint, implemented using the RDKit package.⁷⁶ Morgan fingerprints are calculated based only on 2D molecular structural information, providing a vector representation of substructures for any given molecular input. Using a 1024-bit vector encoding and a maximum of five atoms per substructure, these 2D fingerprints were generated directly from SMILES. To ensure that reactions are correctly indexed, we specifically use atom-mapped SMILES. A key advantage of Morgan fingerprints is that they require very little optimization of their hyper-parameters in order to capture structural changes during the course of a reaction.

The Morgan fingerprint calculation produces a feature vector, \mathbf{d} , for a given input molecular structure; in the case of chemical reactions, we are interested in the impact of the *change* in the molecular structure, which is most closely related to the activation energies being predicted. As such, for each reaction i considered in the training- or test-sets (discussed in the following), the feature vector used as input in the ML models is a difference vector, \mathbf{D}_i , defined as

$$\mathbf{D}_i = \mathbf{d}_i^P - \mathbf{d}_i^R, \quad (1)$$

where \mathbf{d}_i^R and \mathbf{d}_i^P are, respectively, the feature vectors for the reactants or products (Fig. 1).

An important extension to this basic strategy is the inclusion of the reaction energy change as an additional descriptor. As is well-known from the Bronsted–Evans–Polyani (BEP) relation,^{61,62} the activation energy is usually strongly correlated with the reaction energy; in other words, we expect the reaction energy to be a good descriptor for prediction of activation energies, as has also been confirmed by previous investigations.^{58,63–65} Although the focus of this study is to predict activation energies, we note that our ML strategy in the following could, in principle, be applied to predict the reaction energies also, potentially avoiding geometry optimization calculations altogether; however, here, we focus simply on activation energy prediction, assuming that geometry optimization for reaction end-points is relatively straightforward in most anticipated applications.

3. ANN architecture

For this article, we investigated the performance of several ML models for predicting activation energies. In all cases, we estimated activation energies using the difference vector [Eq. (1)] and the *ab initio* reaction energies as inputs. In particular, we considered kernel ridge regression, decision trees such as random forests, Gaussian process regression, and ANNs. Ultimately, we chose to use ANNs due to small improvements in testing errors relative to the other regression methods for this specific application.

ANNs were built using the `scikit-learn` package.⁷⁷ We partitioned the Grambow dataset of 32 730 total reactions (including both forward and backward reactions) into an 80:20 training–test split. This was performed randomly in accordance with k -fold cross validation,⁷⁸ where $k = 5$; a total of three different ANNs were

trained and used to deliver ensemble-averaged predictions of activation energies. This allows us to evaluate the performance of our ML model’s ability to predict out-of-sample activation energies. We optimized the ANN architecture using `scikit-learn`’s `GridSearchCV` function, which identified 200 nodes in two hidden layers as the optimal choice for our ANN model. ANN weights were optimized using the Adam optimizer with an initial learning rate of 0.001 and a Rectified linear unit (ReLU) activation function.

As an aside, we note that the ANN prediction errors in the training- and test-sets are significantly reduced if the reaction energy descriptor is not normalized. Again, we believe that this reflects the underlying strong correlation between reaction energy and activation energy, which is reinforced when normalization is not applied to the input reaction energies. As discussed in the following, the performance of our ML model is evaluated using standard error metrics, such as the MAE, RMSE, and R^2 .

B. Microkinetics simulations

For each of the CRNs considered in the following, we perform ML predictions of activation energies for all relevant chemical reactions defined in the network; we note that the input relative energies of reactants and products are known from prior calculations, as discussed in the following.

Subsequently, we use standard TST to approximate the corresponding reaction rates of all reactions in each studied CRN. Following on from the discussion of ANN training above, it is important to note that the TST rates calculated are not necessarily those that might be calculated using standard *ab initio* methods. In particular, our ANN is trained to predict activation *enthalpies* ΔH^\ddagger , whereas the typical TST rate calculation employs the Gibbs activation energy, ΔG^\ddagger ,

$$k(T) = \frac{k_B T}{h} e^{-\frac{\Delta G^\ddagger}{k_B T}}.$$

Therefore, by training our ANN against a specific set of enthalpies, we inherently ignore the entropic contribution to activation free energies; in addition, the temperature-dependence of this contribution is also obviously ignored. This is, of course, a necessary approximation driven by the availability of appropriate training sets; however, as shown in the following, we find that our ANN is surprisingly robust in predicting activation energies.

With the rates and relative energies in the CRN in hand, we proceed to perform microkinetics simulations using the well-known stochastic simulation algorithm (SSA) developed extensively by Gillespie.^{14,26–28} Here, given the list of current species concentrations and the set of reaction rates, a probabilistic strategy is adopted to select the next reaction to occur, and the simulation time is advanced accordingly; repeating this approach yields a stochastic trajectory of time-dependent species concentrations, with repeated runs, enabling calculation of average time-evolution of concentrations and related uncertainties.

In addition to exploring the impact of ML-predicted activation energies, in this article, we explicitly study the impact of activation energy uncertainties also. These simulations are possible because we are using ANNs to predict activation energies. When using DFT predictions of activation energies, the correct “reference point” to

calculate uncertainties is unknown; however, in the case of ANNs, one can straightforwardly generate approximate uncertainties in predicted activation energies, enabling one to evaluate the impact of errors in microkinetics simulations of CRNs.

To calculate the uncertainties in predicted activation energies, we simply use an ensemble of trained ANNs when making predictions. Each ANN is trained in the same way and has the same structure (as described in the following) but starts from a different initial set of connection weights; weight optimization during training then produces a set of ANNs that all predict the target activation energies but each with a different set of weights.

When determining activation energies for each of the CRNs in the following, we use the ensemble-averaged prediction; similarly, the variance σ^2 in each activation energy can also be evaluated using the ANN ensemble, as can the covariance matrix \mathbf{C} describing correlation between different reactions. With the mean ANN prediction, it is straightforward to perform traditional microkinetics simulations of CRNs to predict the time-evolution of species concentrations; given that prior *ab initio* activation energies are available for each CRN studied in the following, the ANN-based microkinetics can be compared to DFT-based results, enabling evaluation of the impact of ML predictions. Furthermore, in the ANN-based simulations, the uncertainties in activation energies are incorporated here in a straightforward way; in particular, for each CRN, we perform a series of three independent simulations, each using a different set of activation energies for each CRN drawn from a multivariate Gaussian distribution given by

$$P(\mathbf{E}) = N e^{-[(\mathbf{E}-\langle\mathbf{E}\rangle)^T \mathbf{C}^{-1} (\mathbf{E}-\langle\mathbf{E}\rangle)]}$$

Here, \mathbf{E} is the set of activation energies sampled from the multivariate Gaussian distribution $P(\mathbf{E})$, $\langle\mathbf{E}\rangle$ is the set of ANN-predicted mean activation energies, and N is the corresponding normalization factor. Before each SSA simulation, activation energies are sampled from $P(\mathbf{E})$ using a standard Metropolis Monte Carlo procedure. In this way, averaging over barrier heights and accounting for covariance among the different barrier heights, our ANN-based microkinetics simulations offer a first-order account of ANN predictive uncertainty and its impact on microkinetics in CRNs.

III. APPLICATIONS AND DISCUSSION

In this section, we begin by demonstrating the accuracy of activation energy prediction using our ANN architecture. Subsequently, we consider the impact of ANN activation energy prediction (and uncertainties) in three different organic chemistry CRNs, each with different characteristics.

A. ANN prediction accuracy

Figure 2 shows performance metrics for our ANNs; after training, the ensemble-averaged MAE and RMSE for a test-set of 6546 reactions were 2.78 and 3.82 kcal mol⁻¹, respectively. The correlation plot in Fig. 2(a) shows a positive linear regression profile for the activation energies in both the training- and test-sets of reactions ($R^2 = 0.97$). These results clearly demonstrate good correlation between ML- and DFT-calculated activation energies in the training- and test-sets and show no evidence of consistent over- or under-estimation of barriers. Relative to DFT results, the vast

majority of the errors fall into the ± 20 kcal mol⁻¹ error range with most reactions reporting error values of under ± 10 kcal mol⁻¹. Furthermore, in Fig. 2(b), we illustrate how ANN accuracy changes with training-set size. Both the test and training curves converge at around 16 000 training points with MAEs of 2.78 and 2.74 kcal mol⁻¹, respectively. The R^2 accuracy reaches a maximum value at roughly the same point. Our model's overall accuracy and performance are therefore not impacted by adding more training samples beyond about 16 000 reactions.

Overall, our analysis of barrier predictions appears to be comparable to the levels of accuracy obtained in previous studies using a variety of different ML architectures. For example, as noted above, previous investigations have obtained RMSE and MAE prediction errors of around 2–6 kcal mol⁻¹, which is clearly comparable to our results. As such, our trained ANN is broadly representative of typical accessible prediction accuracy; our results in the following assess the impact of this level of accuracy.

B. ANN predictions for microkinetics simulations

Having verified that our ANN predicts activation energies for organic chemical reactions that are comparable to previous studies, we now consider the impact of ANN predictions on microkinetic modeling.

To do so, we consider three different CRNs, which have been previously studied and which were either generated “by hand” or by using automated reaction discovery tools. To be truly useful, ANN predictions of activation energies should ideally be applicable to heterogeneous calculation setups; in other words, the accuracy of ANN predictions should not be tied to any particular reaction discovery method or *ab initio* electronic structure method for energy evaluation. As such, our approach to testing our ANN predictions is to take reaction energetics and molecular structures from previous CRNs generated by using either DFT or coupled-cluster with singles, double and perturbative triples [CCSD(T)]; our simulations in the following will then test whether ANN predictions of activation energies are useful, without requiring complete re-calculation of energies and other properties of all species in a given known CRN.

The CRNs studied here are (i) decomposition of formamide, (ii) aldol reaction between vinyl alcohol (H₂C=CHOH) and formaldehyde (H₂CO), and (iii) unimolecular decomposition of 3-hydroperoxypropanal. These CRNs were chosen because they each have different overall structures and complexity but also have available *ab initio* electronic structure calculations for comparison with ML predictions. First, the formamide network is a simple “linear” CRN with a set of four coupled chemical reactions, leading along two pathways to the same reaction product; furthermore, the set of activation barriers for all reactions is quite comparable in magnitude (Table I, see in the following). The aldol network has a more centralized structure with multiple possible products. Finally, the unimolecular decomposition of 3-hydroperoxypropanal is another example of a CRN with a centralized structure, representing an organic decomposition mechanism starting from a single molecule and leading to 16 possible products; in this case, the activation energies in the CRN are such that one reaction lies about 20 kcal mol⁻¹ lower than the rest of the reactions (Table III).

Using the ANN setup described above, we predicted the activation energies for all relevant reactions in these CRNs. Figure 3 gives

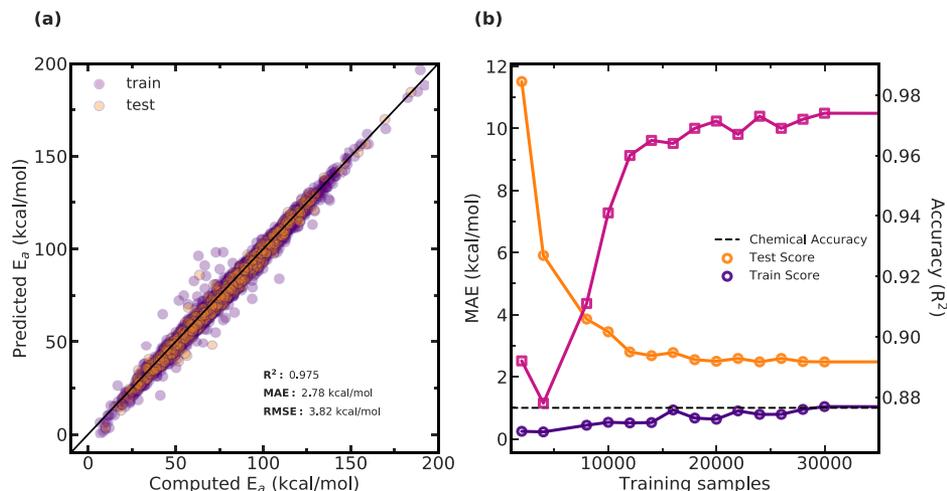


FIG. 2. Trained ANN model using the Morgan fingerprint molecular descriptor. (a) Correlation plot of DFT-calculated and ANN-predicted activation energies. The performance of the fingerprint model for the train (purple circles) and test (orange circles) sets are quantified by key error metrics such as the mean absolute error (MAE), root mean squared error (RMSE), and the coefficient of determination (R^2) (pink square), shown here in (b) the learning curve. Note that the dotted line in the training curve represents the standard 1 kcal mol^{-1} target for chemical accuracy.

TABLE I. CCSD(T) calculations and ANN predictions of activation energies for reactions in the formamide (NH_2CHO) decomposition CRN as well as the standard deviation (σ) of the ensemble average. The value Δ is simply the difference between ANN-predicted and *ab initio* computed activation energies. All values are given in kcal mol^{-1} , and the CCSD(T) literature results appear in Fig. 5 in the article by Nguyen *et al.*⁷²

Reaction index	CCSD(T)	ML	Δ	σ
1	43.90	41.81	-2.09	0.52
2	77.00	84.75	7.75	2.50
3	72.20	61.30	-10.90	0.86
4	38.90	40.71	1.81	1.46

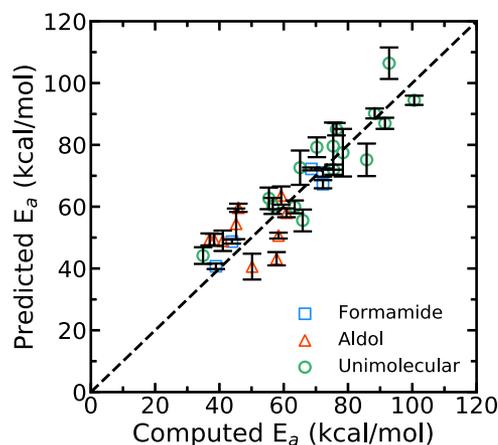


FIG. 3. Computed vs machine-learning predicted barriers for three different sets of organic reactions. The decomposition of formamide⁷² (blue squares), the aldol reaction between vinyl alcohol and formaldehyde⁴³ (red triangles), and the unimolecular decomposition of 3-hydroperoxypropanal³² (green circles). The vertical error bars represent the standard deviations of the predicted barriers.

an overview of the ANN activation energy predictions for the three different CRNs considered here; in addition, the uncertainties on the ANN predictions are also shown. In each case, we compare the ANN predictions to the barrier heights obtained in previous *ab initio* calculations [either DFT or CCSD(T)]; a list of the calculated barrier heights, uncertainties, and reaction identities is also included in the [supplementary material](#).

First, from Fig. 3, it is clear that the ANN approximation of activation energies is reasonably accurate. There is clearly good agreement between the activation energies predicted by *ab initio* calculations in previous work and in the ANN predictions; the calculated coefficient of determination is $R^2 = 0.91$. Furthermore, we find that the level of agreement is maintained across a broad range of chemically relevant activation barriers, namely, from about 35 to $100 \text{ kcal mol}^{-1}$. Perhaps most surprisingly, the level of agreement between ANN predictions and *ab initio* results is good despite the fact that the reaction energies that were used to generate the ANN inputs for these reactions were different from the method used in the original training data; for example, in the formamide CRN, the initial reaction energetics were calculated using CCSD(T), not the same DFT B97-D3/def2-mSVP level-of-theory as used for the training set.

1. Analysis of reaction ordering

Before proceeding to discuss the microkinetics simulation results, it is useful to investigate whether the use of ANN predictions of activation barriers actually changes the preference for different reactions in each CRN. In the best-case, one might hope that the barrier-ordered list of reactions is the same in both *ab initio* calculations and ANN prediction calculations; in the worst-case, it is possible that changes in activation energies could lead to switching of preferred pathways through a CRN.

Figure 4 shows the activation energies for each CRN plotted by the reaction index; these plots enable identification of reactions for which “switches” in barrier heights between *ab initio* and ANN results are evident. For the formamide CRN, it is clear that the ANN predictions closely follow the trend in the activation energies for the four predicted reactions; for example, reaction 2 is found to have the largest barrier in both ANN and *ab initio* predictions, whereas

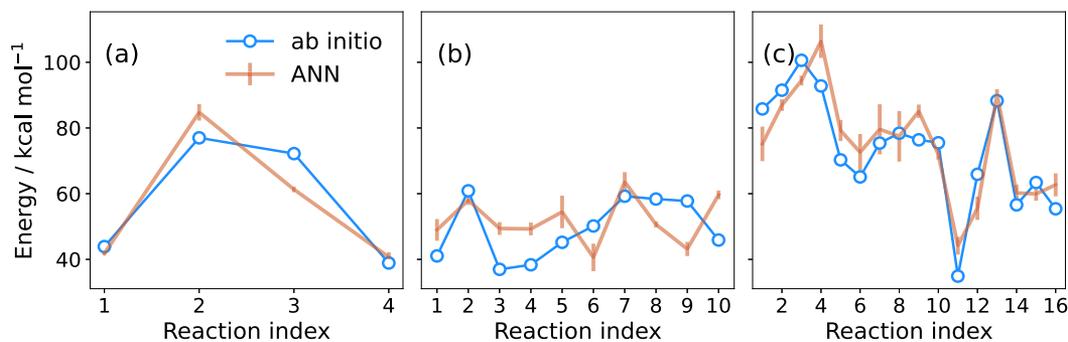


FIG. 4. Comparison of *ab initio* and ANN activation energy predictions for (a) formamide CRN, (b) aldol CRN, and (c) CRN describing unimolecular decomposition of 3-hydroperoxypropanal. Ensemble standard deviations for ANN predictions are illustrated; the lines are intended as a guide for the eye only.

reactions 1 and 4 are much lower (and predicted to be essentially the same in both ANN and *ab initio* calculations). In contrast, the barrier heights for the aldol reaction show some significant differences; for example, the DFT calculations predict that reaction 3 has the lowest activation energy, whereas the ANN predicts that reaction 6 has the lowest activation energy. It is worth noting, however, that the range of activation energies spanned by this CRN is quite small, and all predictions by the ANN sit within the expected range for this CRN. Finally, the ANN predictions for the unimolecular decomposition reaction are found to be very good indeed. In particular, reaction 11 is clearly the lowest-energy barrier in both DFT and ANN predictions with the trend in the remaining higher barriers being well-reproduced by the ANN.

From Fig. 4, we might therefore expect that the ANN-based kinetics simulations of the formamide and unimolecular decomposition CRNs would be broadly in agreement with the kinetics generated by the *ab initio* data, whereas the simulations of the aldol CRN might be expected to be quite different between these two different activation energy sets. This is indeed found to be the case, but the kinetics results in the following provide much more insight than what might be evident from Fig. 4 alone.

2. Analysis of ANN ensemble uncertainties

As noted above, we use an ensemble of ANNs to predict activation energies; this also enables evaluation of the covariance matrix, which is used in some of the microkinetic simulations in the following. Broadly, one might expect the standard deviations in the ANN-predicted activation energies to correlate with the errors between the average ANN predictions and the *ab initio* calculations; in other words, when activation energies are poorly reproduced, one would hope that the ensemble of ANNs recognize this in a larger uncertainty among the predictions.

Figure 5 shows the correlation between the standard deviations σ from the ANN ensemble and the absolute magnitude of the deviation $\Delta = E_{\text{ML}} - E_{\text{ab-initio}}$. For both the formamide and unimolecular decomposition CRNs, we see a broad correlation between these properties; this is most evident for the unimolecular decomposition CRN, where we typically find that reactions with small uncertainties are accurately predicted. However, we do find that the ANN ensemble uncertainties tend to underestimate the true error relative to the *ab initio* results. For the aldol CRN, the correlation is less clear-cut;

it seems that the ANN ensemble uncertainties are generally much lower than the true errors in the predicted barriers, consistent with the discussion of Fig. 4. The reactions encompassed by the aldol CRN, and the system-size, seem comparable to the formamide or unimolecular decomposition CRNs, so one might expect similar correlation; however, we also note that the ANN predictions for each of the CRNs are being compared against different *ab initio* methods, which complicates the comparison here. However, for the purposes of this article, we simply note that there is some broad correlation between uncertainties and true errors although the ANN uncertainties generally appear to be underestimated. It is also worth noting that given the approach taken to training an ANN ensemble here (i.e., same structures but different weights), it is perhaps not surprising that the ANN uncertainties are underestimated; of course, methods such as GPR would be expected to give more reliable error estimates, but we have already noted above that the overall performance of GPR for the particular regression task studied here was not as good as ANN.

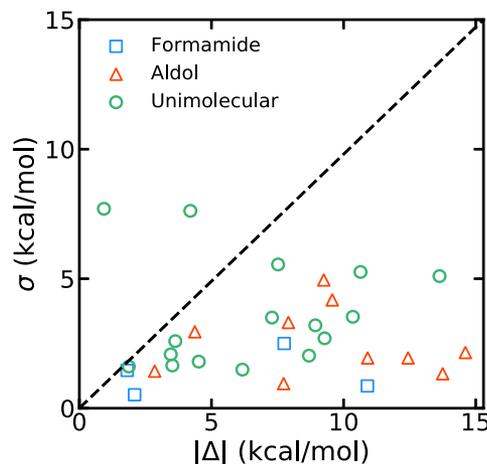


FIG. 5. Comparison of ANN-predicted uncertainties (σ) with errors between ANN-predicted activation energies and *ab initio* activation energies (Δ). Results are shown for each of the three CRNs studied here.

C. Kinetics simulations

In the following, we consider the results of microkinetic simulations performed for each of the three CRNs with activation energies determined through three different routes, namely, (i) *ab initio* calculations, (ii) ANN prediction, and (iii) ANN prediction with account of activation energy covariance.

1. Formamide reaction network

First, we consider the CRN for formamide (NH_2CHO) decomposition; as noted previously, this molecule has been extensively studied as part of prebiotic pathways toward complex biomolecules. Here, the molecular geometries were taken from a previous study that employed CCSD(T) calculations to obtain relative energies of intermediates and TSs for a sequence of decomposition reactions. We particularly focus here on decomposition of formamide into $\text{HNCO} + \text{H}_2$.

Despite the relative simplicity of this system, these simulations immediately reveal key problems with ANN strategies of the type employed here. The originally studied CCSD(T) reaction network contains a total of seven reaction barriers; however, a number of these are barriers to *conformational* isomerizations in which the bonding of intermediate structures does not change but the relative torsion angles do. In using connectivity-based fingerprints, such as the Morgan fingerprint scheme employed here, the fingerprints calculated for different conformers are equivalent such that the ANN cannot be used to accurately predict the associated activation energies. This is a significant and well-known disadvantage of descriptors based solely on connectivity information, meaning that activation energies for conformational interconversions cannot be reliably predicted by ANNs trained using these fingerprints alone. We note that in developing the ANN described above, we have tried to use alternative fingerprints, such as the well-known smooth overlap of atomic position (SOAP) descriptors;⁷⁹ such descriptors do capture the differences between different three-dimensional conformations, but we have found so far that their predictive performance in activation energy determination was not as good as the Morgan fingerprints used here. Further optimization of this approach might be expected to improve performance but sits outside the scope of this article.

To account for this conformational problem, the CRN studied here comprises four (rather than seven) reactions, which lead to $\text{H}_2 + \text{HNCO}$ along two double-step mechanisms from NH_2CHO . The reactions corresponding to conformational changes have been removed; in this case, this is somewhat further justified by the fact that these reactions have generally lower barriers than those related to chemical reactions (bonding changes) such that one can make the assumption that the multiple conformations equilibrate within themselves and are replaced by a single effective conformer.

To predict activation energies using our trained ANN, the molecular geometries for reactants and products for each reaction in the CRN were obtained from previous work using CCSD(T) calculations; these were subsequently used as input for our ANN. As noted above, our trained ANN also requires input of the reaction energy for each target reaction for prediction; these values were taken directly from the previous investigation.⁷²

The ANN-predicted activation energies for the series of formamide reactions considered here are shown in Table I. As also

evident from Fig. 3, the agreement between the ML-predicted barriers and those obtained by CCSD(T) is very good; furthermore, the calculated (ensemble-averaged) standard deviation in the ML-predicted barrier heights is also quite low, found to be less than or equal to $2.5 \text{ kcal mol}^{-1}$ in all cases. These results indicate that the ANN-predicted barriers are surprisingly robust to changes in the electronic structure method; our ANN was trained using reaction energetics obtained using DFT, but we find here that the predictions made using CCSD(T) reaction energies as input are just as good. This is a useful observation, suggesting transferability across different electronic structure methods for input.

Figure 6 shows the time-dependence of the concentration of the initial formamide reactant [Fig. 6(a)] and the intermediate structure NHCHOH [Fig. 6(b)] during SSA simulations at short times (up to $2 \times 10^{-4} \text{ s}$). Here, the initial concentration of formamide was (arbitrarily) chosen as 1.33 mol dm^{-3} . Different SSA simulations were performed using (i) the original CCSD(T)-calculated barriers, (ii) ANN-predicted barriers with no account for uncertainties (i.e., simply using the averaged predicted barrier), and (iii) ANN-predicted barriers with covariance accounted for as described in Sec. II. In total, three independent SSA simulations were performed in each case; the results of Fig. 6 show the averaged concentrations

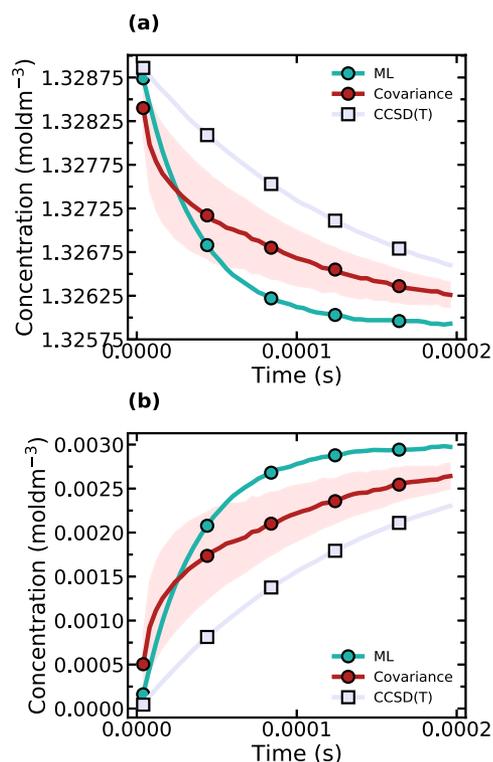


FIG. 6. Time-dependent concentrations for formamide (NH_2CHO) decomposition reactions at $T = 800 \text{ K}$. Activation energies were determined using CCSD(T) calculations (purple squares), ML (green circles), and ML with covariance (red circles). Results are shown for the populations of (a) NH_2CHO and (b) NHCHOH .

and the associated uncertainties. It is important to note that the described microkinetic simulation setup reflects an idealized system, predominantly chosen for computational convenience rather than representing real-world experimental conditions. However, for the purposes of this article, it is less relevant as it does not impact the comparisons made between DFT computed and ML error-qualified kinetics.

In keeping with the accuracy of the ML-predicted activation energies in Fig. 3, the *ab initio* and ML time–concentration plots are in good agreement, especially when considering the small concentration changes involved. For example, the timescales associated with the decay of the initial reactant generally agree well across the different methods. Additionally, the impact of uncertainties appears to have a relatively small influence on the overall kinetics; this is perhaps unsurprising, given that errors in the ANN-predicted barriers (Table I) are less than $2.5 \text{ kcal mol}^{-1}$.

However, looking at the results of Table I, it is clear that the good agreement between the microkinetic properties predicted by *ab initio* calculations and ML predictions observed in Fig. 6 masks some of the underlying prediction errors. In particular, the errors in the ML-predicted activation energies for reactions 2 and 3 are around 8 and 10 kcal mol^{-1} , respectively; these are significant errors in the ANN.

To explain this discrepancy, we should examine the overall ranking of reactions within this scenario (Fig. 4). In this case, we find that the overall ordering of activation energy magnitudes for the different reactions remains unchanged with reaction 2 being predicted to have the highest barrier and reaction 4 is predicted to have the lowest barrier, thereby maintaining the overall kinetics observed in this reaction. Further consideration of the flux through different reactions reveals that the formamide network exhibits a preference for progressing via reactions 1 and 2 (as seen in Table I) to generate the final product of $\text{H}_2 + \text{HNCO}$. Therefore, the short-time kinetics is in good agreement between the different methods predominantly because of the good agreement in the lowest-energy barrier, reaction 1 (Table I).

It is worth noting that although there are clear overall similarities in the qualitative behavior for the standard ML-based SSA simulations and those that incorporate uncertainty through the covariance matrix, there are also some differences in the quantitative behavior. We attribute this to the simulation approach taken here. In particular, for the simulations employing uncertainty, we are sampling activation energies from a multivariate Gaussian distribution and subsequently using these activation energies to approximate reaction rates via TST. In other words, we are calculating a function of the sampled activation energies; as such, given that the average value of a function sampled over an input distribution is not necessarily the same as the function calculated for the average input [i.e., $\langle f(x) \rangle \neq f(\langle x \rangle)$, in general], then it is not surprising that some differences are apparent here.

In summary, the findings for the formamide are encouraging; the predicted barriers given by ANN are in good agreement with those obtained from CCSD(T) calculations, and both the short-time and long-time kinetics of the CRN are in broadly good agreement also. These findings will inevitably be influenced by the size of the CRN and should therefore be examined more thoroughly using larger, heterogeneous reaction networks. To that end, Secs. III C 2 and III C 3 study more complex CRNs.

2. Aldol reactions

In the second CRN considered here, focusing on the aldol reaction, molecular geometries for reactants and products were obtained from a previous DFT investigation. Here, we extracted DFT-optimized geometries and DFT-calculated reaction energies for a set of ten reactions; subsequent predictions by our trained ANN using the DFT energetics available from previous work then resulted in the set of activation energies shown in Fig. 3 and in Table II.

As discussed above, there is generally quite good agreement between the ANN activation energies and those obtained previously by DFT studies. However, it is noticeable that there are some quite significant activation energy differences of up to 14 kcal mol^{-1} ; as we show in the following, in contrast to the formamide CRN (where the agreement between *ab initio* and ML-based microkinetics was quite good), these errors can lead to significant differences in the microkinetics predictions.

After activation energy prediction, three different sets of SSA simulations were performed (using ML predictions, ML predictions with sampling over covariance, and DFT activation energies), as described above. Here, the initial concentrations of both vinyl alcohol ($\text{H}_2\text{C}=\text{CH}-\text{OH}$) and formaldehyde (CH_2O) were 1.33 mol dm^{-3} with all other species in the CRN beginning with zero concentration. All SSA simulations were performed for $t = 0.5 \text{ s}$, and results were averaged over three independent simulations.

Figure 7 shows the time-dependent concentrations of the reactants [Figs. 7(a) and 7(b)] as well as two different products [Figs. 7(c) and 7(d)] from SSA simulations of the aldol CRN. The species in (c) and (d) correspond to the products of reactions 2 and 7, respectively (Table II).

Again, we find quite good agreement between ML and DFT-based kinetic predictions with similar kinetic timescales associated with the decay or growth of reactants and products, respectively. However, in contrast to the previous formamide CRN, we observe a greater degree of error when covariance sampling over activation energies is used in the ML predictions; in particular, the populations observed after 0.5 s simulation time are quite different when covariance sampling is used in the ML-based SSA simulations. This is a direct result of the larger observed variances observed for this CRN

TABLE II. A DFT calculation and ML prediction of Aldol reaction activation energies as well as the standard deviation (σ) of the estimates. The value Δ is simply the difference between *ab initio* and ANN-predicted activation energies. All values are in kcal mol^{-1} , and the DFT literature results are from Fig. 5 in Maeda *et al.*⁴³

Reaction index	DFT	ML	Δ	σ
1	41.06	48.97	7.91	3.31
2	60.87	58.02	-2.85	1.43
3	36.97	49.41	12.44	1.94
4	38.36	49.27	10.91	1.94
5	45.20	54.45	9.25	4.95
6	50.17	40.60	-9.57	4.18
7	59.23	63.60	4.37	2.95
8	58.37	50.64	-7.73	0.95
9	57.77	43.16	-14.61	2.15
10	45.89	59.64	13.75	1.33

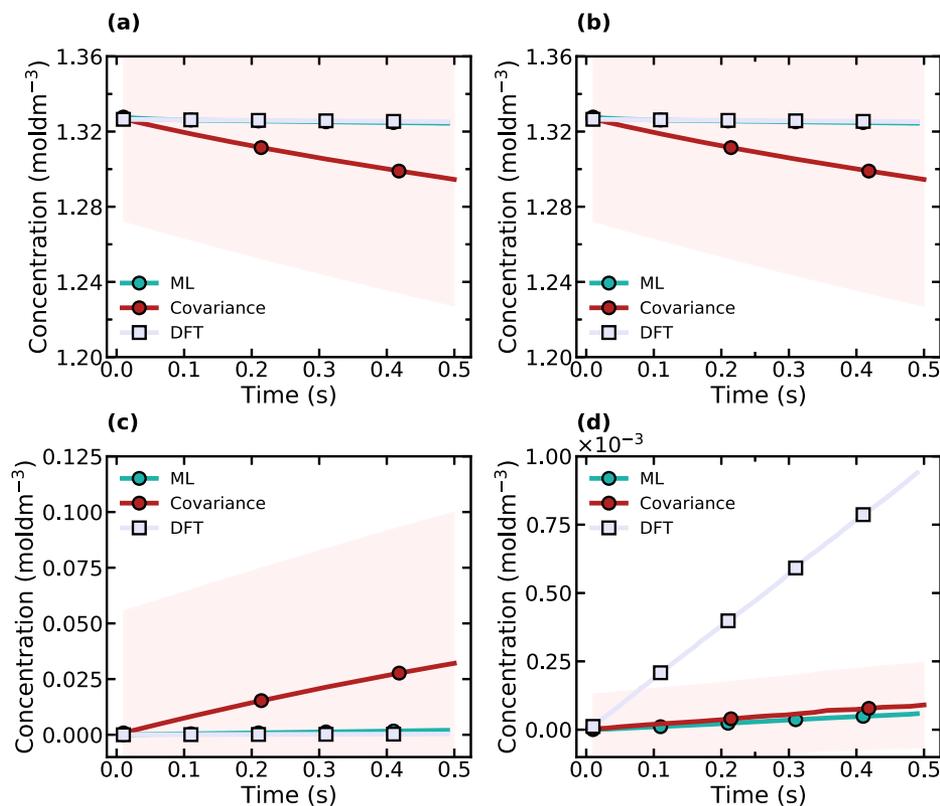


FIG. 7. Time-dependent concentrations for aldol reactions at $T = 800$ K. Activation energies were determined using DFT calculations at the B3LYP/6-31G level (purple squares), ML (green circles), and ML with covariance (red circles). Results are shown for the populations of (a) vinyl alcohol ($\text{H}_2\text{C}=\text{CH}-\text{OH}$), (b) formaldehyde (H_2CO), (c) product species 4, and (d) product species 9 (see the [supplementary material](#)).

(Table II), where we find standard deviations ranging from 0.95 to 4.95.

Perhaps most significantly, we find the kinetics of this CRN are strongly impacted by the ML predictions but not necessarily in a straightforward manner; this is a consequence of the coupled nature of reactions in the CRN. For example, reaction 3 has a lower barrier in *ab initio* calculations than predicted by ANN; however, we find that the populations predicted in the DFT-based and ML-based kinetics simulations for this structure are quite similar with an equilibrium between forward and backward reactions quickly established. A similar observation is found for reaction 2 (forming product species 4; Fig. 7), but in this case, the simulations including covariance are found to give larger concentrations than other methods. Similarly, the concentration of product species 9 [Fig. 7(c)], formed by reaction 7, is quite different in the DFT simulations compared to other methods; we note that species 9 is the thermodynamically stable product in this CRN such that the species concentrations in Fig. 7 reflect the difference in activation energies with less influence from the backward reaction.

Overall, the aldol CRN demonstrates more significant differences between the ML and DFT kinetics modeling, predominantly as a result of differences reflected in Fig. 4; however, we note that this picture is complicated by the backward reactions in the CRN also. Despite this, we note that the order-of-magnitude difference in the

species concentrations in Fig. 7 is broadly in agreement (although, again, this is not universally guaranteed).

3. Unimolecular decomposition of 3-hydroperoxypropanal

The final CRN studied here was the unimolecular decomposition of 3-hydroperoxypropanal, a CRN which had been previously generated by simulations using KinBot,⁸⁰ an automatic PES exploration program. Using molecular geometries obtained from previous investigations of this CRN, ANN-predicted activation energies were obtained using previous DFT reaction energetics as input. In total, the CRN comprised 16 selected reactions, as described in the [supplementary material](#).

The original DFT and ANN-predicted activation energies are shown in Table III. As expected from Fig. 3, the ANN-predicted barriers are in reasonably good agreement with those obtained from previous DFT calculations; again, we note that this underlines the robust nature of the ANN predictions, given the fact that the input DFT data were obtained using B3LYP/6-31+G*, whereas the ANN training data were calculated at the B97-D3/def2-mSVP level. However, it is noticeable that some of the barrier predictions given by the ANN differ from the DFT results by up to 13 kcal mol⁻¹.

Following ANN prediction, three sets of SSA simulations were performed as described above. The initial concentration of 3-hydroperoxypropanal was 1.33 mol dm⁻³ with all other species

TABLE III. DFT calculations and ANN predictions of activation energies for selected reactions in CRN describing unimolecular decomposition of 3-hydroperoxypropanal. Here, σ is the standard deviation in the ensemble ANN predictions, and Δ is simply the difference between *ab initio* and ANN-predicted activation energies. All values are in kcal mol⁻¹, and the DFT literature results are from Figs. 5–7 in the supplementary material of Grambow *et al.*³²

Reaction index	DFT	ML	Δ	σ
1	85.80	75.16	-10.64	5.26
2	91.50	86.98	-4.52	1.80
3	100.60	94.43	-6.17	1.49
4	92.80	106.43	13.63	5.10
5	70.30	79.23	8.93	3.20
6	65.10	72.62	7.52	5.55
7	75.40	79.60	4.20	7.62
8	78.40	77.47	-0.93	7.70
9	76.40	85.09	8.69	2.03
10	75.50	71.98	-3.52	1.65
11	34.90	44.18	9.28	2.70
12	65.90	55.55	-10.35	3.53
13	88.30	90.17	1.87	1.61
14	56.60	60.23	3.63	2.59
15	63.40	59.94	-3.46	2.08
16	55.40	62.69	7.29	3.50

having zero initial concentrations. The SSA simulations were propagated for a maximum of 0.5 s, and the results were averaged over five independent calculations.

Figure 8 shows the time-dependence of the concentrations of 3-hydroperoxypropanal [i.e., the reactant species, shown here in Figs. 8(a) and 8(b)] and the product, 1,2-dioxolan-3-ol [Figs. 8(c) and 8(d)]; we note that the reaction forming 1,2-dioxolan-3-ol from 3-hydroperoxypropanal is the reaction with the lowest activation barrier predicted by both DFT and our ANN (i.e., reaction 11 in Table III). According to the original literature study,³² this was also found to be the major reaction pathway, confirmed by several transition-state finding methods.

Most importantly, in broad contrast to the simulations of formamide and aldol CRNs, it is clear that there is a very large difference in the CRN kinetics predicted using ANN and DFT activation energies. In particular, in simulations using DFT-predicted activation energies [Figs. 8(b) and 8(d)], the reaction effectively proceeds to completion after 10⁻³ s with all of the initial 3-hydroperoxypropanal converted into 1,2-dioxolan-3-ol. In contrast, the kinetics simulations performed using ANN-predicted barriers exhibit a *much* longer kinetic timescale; here, the reaction is (mostly) complete after about 0.3 s. However, we note that in both DFT- and ANN-based kinetics simulations, there is no difference in the final reaction products or the concentration of other product species (which are all essentially zero). As such, this is a clear example of the

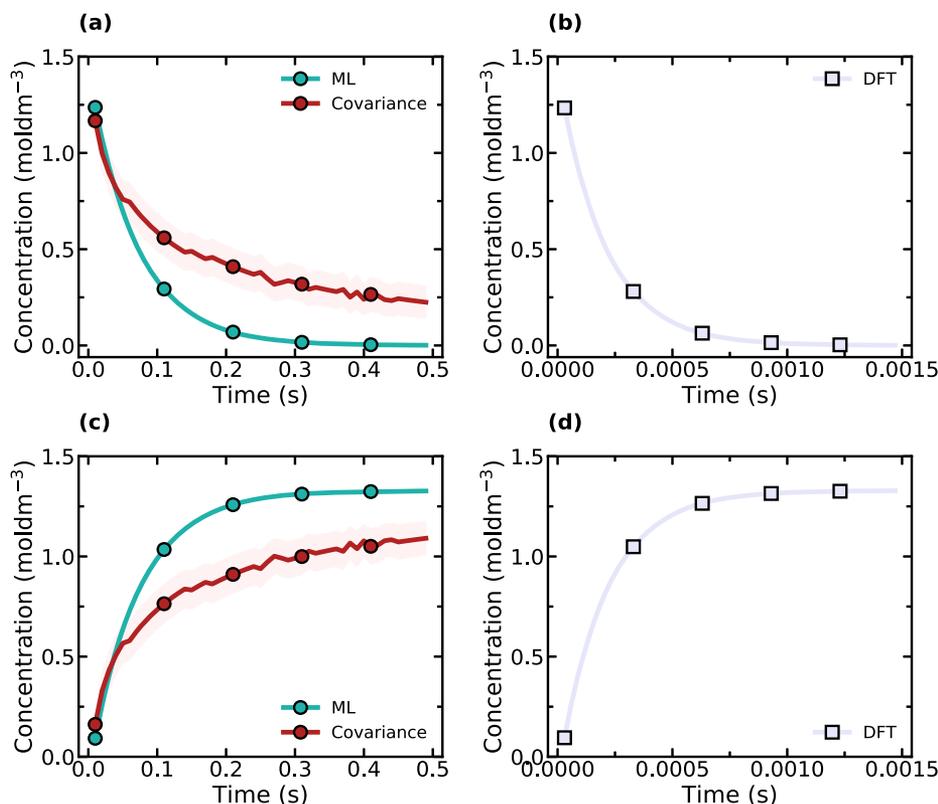


FIG. 8. Time-dependent concentrations for the unimolecular decomposition of 3-hydroperoxypropanal at $T = 800$ K. Activation energies were determined using DFT calculations at the B3LYP/6-31+G* level (purple squares), ML (green circles), and ML with covariance (red circles). Results are shown for the populations of (a) and (b) 3-hydroperoxypropanal using ANN and DFT-calculated barriers, respectively, for (c) and (d) unimolecular reaction 11 using ANN and DFT-calculated barriers, respectively.

ANN-predicted CRN exhibiting strongly different timescales to that observed in DFT-based simulations.

To explain this stark difference, we considered the predicted barriers in more detail. The main reaction product being formed in this network is 1,2-dioxolan-3-ol, which is generated by 3-hydroperoxypropanal decomposing *via* reaction 11 in Table III, identified in literature as being the first step in the Korcek mechanism.⁸¹

It is clear that reaction 11 has a much lower activation energy by 20–30 kcal mol⁻¹ relative to all other reactions in the CRNs. As such and based on the SSA simulation results, it is clear that the CRN is dominated by this single reaction such that the timescale associated with Fig. 8 can be related directly to the rate of this reaction alone.

On the basis of TST, the relative difference in the rate of the important reaction 11 when using either DFT or ANN activation energies should be given by

$$k_{\text{rel}} = e^{-\frac{\Delta\Delta G^\ddagger}{k_B T}},$$

where $\Delta\Delta G^\ddagger = G_{\text{DFT}}^\ddagger - G_{\text{ANN}}^\ddagger$ is the relative activation energy difference arising from predictions given by DFT and ANN. Using the activation energies from Table III, we find $k_{\text{rel}} = 3 \times 10^{-3}$. This is a significant difference in timescale, which correlates strongly with the observed difference in the kinetic timescales observed in the SSA simulations of Fig. 8.

D. Summary

The determination of activation barriers using standard quantum chemistry tools, such as TS-finding methods, is a laborious task—particularly if one is interested in large-scale CRNs. Here, we have studied an alternative data driven scheme using ML to predict activation energies for reactions at a fraction of the cost offered by most *ab initio* calculations. To investigate the impact of ML-predicted activation energies, we have performed a series of microkinetics simulations of CRNs in which the activation energies for the component reactions were obtained by both ANN predictions and *ab initio* calculations; our ANN was trained using a training set comprising ~32 000 organic reactions and was found to exhibit RMSE/MAE prediction errors of 3.82 and 2.78 kcal/mol, respectively. Importantly, this level of performance is directly comparable to many other recent investigations into predicting activation energies using ML schemes.

Our microkinetics simulations investigated three different CRNs with different underlying structure and activation-barrier characteristics. ML-driven simulations of the formamide CRN were able to generally reproduce the observed kinetics for the same CRN when employing *ab initio*-based activation energies; however, for the aldol and unimolecular decomposition CRNs, we found clear differences between ANN and *ab initio* activation energy predictions. In the case of the aldol network, as shown in Fig. 4, we found that the ANN can predict a different ordering of activation energies such that different reactions become more favorable compared to the kinetics simulations performed with *ab initio* barriers. The results for kinetics simulations of the unimolecular decomposition of 3-hydroperoxypropanal also exhibited significant differences between CRNs generated with either ANN-predicted or *ab initio*-calculated activation energies with the *ab initio* results

leading to kinetic timescales that were a few orders-of-magnitude faster than the ANN-generated CRN. We have tracked this difference to the characteristics of the underlying CRN; in the unimolecular decomposition of 3-hydroperoxypropanal, even though the CRN used here comprises 16 different reactions, only one of these reactions has any bearing on the observed kinetics. As such, any variation in this one barrier caused by the inherent approximation and uncertainty in the ANN will be observed to have a significant impact on the emergent kinetics; this is exactly what we observe with the ANN *over-predicting* this key activation energy and so leading to a significantly longer timescale in the predicted kinetics.

Together, our kinetics simulations highlight five key points:

1. First, ANNs, even the relatively straightforward implementation employed here, can often be used to reliably predict activation energies to a sufficient level of accuracy to enable qualitatively correct kinetic modeling of complex CRNs; this is a highly encouraging result if one is interested in using ANNs as part of automated workflows to construct and simulate large-scale CRNs.
2. Second, however, one must be careful in using ANN predictions in systems that are dominated by a few key reactions; as we have shown in the case of the unimolecular decomposition of 3-hydroperoxypropanal, CRNs that are dominated by a small number of reactions will (in this case, one key reaction) exhibit a higher degree of sensitivity to uncertainties associated with ANN predictions.
3. We have found in our simulations that errors in ANN predictions can significantly alter the underlying kinetic mechanism exhibited by the CRN; this was most evident in the case of the aldol CRN although we note that the remaining two CRNs did not exhibit this property.
4. Fourth, we noted that the choice of descriptors can have a clear impact on the capability of an ANN in activation energy prediction to the extent that some reactions might not even be accessible to predictions; this is an obvious consequence of choosing to use connectivity-based fingerprints in this article such that conformation transitions in the formamide CRN could not be captured but suggests that careful consideration should be given to descriptors and whether they are relevant to the system at hand.
5. Fifth, and finally, we found that the uncertainties predicted by the ANN ensemble are broadly in line with the errors between the ANN activation energies and the true *ab initio* activation energies; however, we also found that the ANN uncertainties are typically much smaller than the true error magnitude, and performance is not equal across different CRNs. As such, based on our experience here, the ANN ensemble uncertainties typically act as a lower bound on the true errors. However, this inaccuracy in uncertainty predictions has been found to be typical of ANN studies, as discussed in previous literature;^{82–84} methods such as using a bootstrap ensemble of ANNs may help in quantifying better estimates of uncertainty in such cases.⁸³ In addition, we note that other sources of uncertainty, for example, relating to optimization of the ANN neural network architecture, would not be expected to be accurately represented by our current approach; furthermore, inherent uncertainties associated with construction of

the training set (such as consideration of conformational flexibility) itself will also undoubtedly play a role. In short, while our current approach represents one standard route to incorporating uncertainty, it is clear this is a topic that extends beyond the scope of the “zero-order” strategy tested here.

Together, these observations highlight the enormous potential for ANNs in modeling complex CRNs but also serve to highlight simple situations where one should take care in practical applications.

IV. CONCLUSIONS

The last few years have seen a dramatic increase in the number of different methods for automated curation and characterization of complex CRNs as well as datasets and ML strategies for predicting activation energies for the reactions in CRNs. The ultimate goal of much of this work is to simplify the creation of CRN models for kinetics simulations; ML is set to transform this field by dramatically simplifying this workflow.

In this article, we have investigated the extent to which current ML strategies for activation energy prediction are “good enough” to be useful in microkinetics simulations of CRNs. Here, we have trained an ANN using a readily available organic chemistry dataset containing reactant/product structures, reaction energetics, and activation barriers; our resulting ANN model is comparable in its accuracy to previous similar ML studies. Importantly, our ANN approach uses standard “out-of-the-box” fingerprints and ANN architecture; as such, it represents a strategy that should be widely applicable (provided sufficient training data are available). Using ANN-predicted activation energies and uncertainties, we subsequently performed microkinetics simulations of three different CRNs with different characteristics and demonstrated the influence that over-prediction of activation barriers of competing reactions can have on overall observed kinetics.

This article has also highlighted a number of requirements and areas for further work in order to continue the growth in application of ML to studies of complex CRNs. The availability of reliable and consistent training data spanning a wide range of chemical reaction spaces seems to be desirable; datasets such as that curated by Grambow and co-workers (and employed here) are an enormously useful resource to the wider community. On a related note, the development of improved automated TS-location methods would make a significant impact on data curation as well. Finally, we note that experimental kinetic data are also highly desirable in CRN studies, providing a further possible route to benchmarking and validation of auto-generated computational CRNs. However, the results demonstrated here show that ML-based predictions of activation energies can be qualitatively correct in analyzing relative trends across different reaction sets; as such, these fast approximation schemes are increasingly well-suited to function as part of high-throughput CRN generation tools in which ML predictions can be used to quickly limit the growth of a CRN during reaction discovery simulations. We are currently integrating this strategy within our graph-based simulations in this domain. In summary, however, it seems clear that ML predictions of chemical reaction characteristics will continue to improve in the immediate future, offering new routes to studying large-scale, complex CRNs.

SUPPLEMENTARY MATERIAL

The [supplementary material](#) contains further details of reactions used in constructing the three CRNs studied here as well as details of ANN predictions for these reactions.

ACKNOWLEDGMENTS

The authors acknowledge support from the Engineering and Physical Sciences Research Council, UK, through Grant Nos. EP/S022848/1 and EP/R020477/1. The authors also acknowledge the Scientific Computing Research Technology Platform at the University of Warwick, UK, for provision of high-performance computing facilities.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

I. Ismail: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **C. Robertson:** Conceptualization (equal); Investigation (equal); Methodology (equal); Software (equal); Supervision (equal); Writing – review and editing (equal). **S. Habershon:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Supervision (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal).

DATA AVAILABILITY

Data from Figs. 2–8 are available at wrap.warwick.ac.uk/166748.

REFERENCES

- ¹C. A. Class, M. Liu, A. G. Vandeputte, and W. H. Green, “Automatic mechanism generation for pyrolysis of di-*tert*-butyl sulfide,” *Phys. Chem. Chem. Phys.* **18**, 21651–21658 (2016).
- ²G. N. Simm and M. Reiher, “Context-driven exploration of complex chemical reaction networks,” *J. Chem. Theory Comput.* **13**, 6108–6119 (2017).
- ³Y. Kim, J. W. Kim, Z. Kim, and W. Y. Kim, “Efficient prediction of reaction paths through molecular graph and reaction network analysis,” *Chem. Sci.* **9**, 825–835 (2018).
- ⁴G. N. Simm, A. C. Vaucher, and M. Reiher, “Exploration of reaction pathways and chemical transformation networks,” *J. Phys. Chem. A* **123**, 385–399 (2019).
- ⁵F. Pietrucci and A. M. Saitta, “Formamide reaction network in gas phase and solution via a unified theoretical approach: Toward a reconciliation of different prebiotic scenarios,” *Proc. Natl. Acad. Sci. U. S. A.* **112**, 15030 (2015).
- ⁶J. Proppe and M. Reiher, “Mechanism deduction from noisy chemical reaction networks,” *J. Chem. Theory Comput.* **15**, 357–370 (2019).
- ⁷A. L. Dewyer, A. J. Argüelles, and P. M. Zimmerman, “Methods for exploring reaction space in molecular systems,” *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**, e1354 (2018).
- ⁸C. Robertson, I. Ismail, and S. Habershon, “Traversing dense networks of elementary chemical reactions to predict minimum-energy reaction mechanisms,” *ChemSystemsChem* **2**, e1900047 (2020).

- ⁹C. Robertson, R. Hyland, A. J. D. Lacey, S. Havens, and S. Habershon, "Identifying barrierless mechanisms for benzene formation in the interstellar medium using permutational invariant reaction discovery," *J. Chem. Theory Comput.* **17**, 2307–2322 (2021).
- ¹⁰I. Ismail, H. B. V. A. Stuttaford-Fowler, C. Ochan Ashok, C. Robertson, and S. Habershon, "Automatic proposal of multistep reaction mechanisms using a graph-driven search," *J. Phys. Chem. A* **123**, 3407–3417 (2019).
- ¹¹C. Robertson and S. Habershon, "Fast screening of homogeneous catalysis mechanisms using graph-driven searches and approximate quantum chemistry," *Catal. Sci. Technol.* **9**, 6357–6369 (2019).
- ¹²S. Habershon, "Automated prediction of catalytic mechanism and rate law using graph-based reaction path sampling," *J. Chem. Theory Comput.* **12**, 1786–1798 (2016).
- ¹³L. E. Rush, P. G. Pringle, and J. N. Harvey, "Computational kinetics of cobalt-catalyzed hydroformylation," *Angew. Chem., Int. Ed.* **53**, 8672–8676 (2014).
- ¹⁴D. T. Gillespie, A. Hellander, and L. R. Petzold, "Perspective: Stochastic algorithms for chemical kinetics," *J. Chem. Phys.* **138**, 170901 (2013).
- ¹⁵J. A. Miller and S. J. Klippenstein, "Master equation methods in gas phase chemical kinetics," *J. Phys. Chem. A* **110**, 10528–10544 (2006).
- ¹⁶C. J. Pope and J. A. Miller, "Exploring old and new benzene formation pathways in low-pressure premixed flames of aliphatic fuels," *Proc. Combust. Inst.* **28**, 1519–1527 (2000).
- ¹⁷K. J. Laidler, *Chemical Kinetics*, 3rd ed. (Harper Collins, New York, 1987).
- ¹⁸G. R. Wittreich, K. Alexopoulos, and D. G. Vlachos, "Microkinetic modeling of surface catalysis," in *Handbook of Materials Modeling: Applications: Current and Emerging Materials*, edited by W. Andreoni and S. Yip (Springer International Publishing, Cham, 2020), pp. 1377–1404.
- ¹⁹S. Blau, H. Patel, E. Spotte-Smith, X. Xie, S. Dwaraknath, and K. Persson, "A chemically consistent graph architecture for massive reaction networks applied to solid-electrolyte interphase formation," *Chem. Sci.* **12**(13), 4931–4939 (2021).
- ²⁰D. G. Truhlar, B. C. Garrett, and S. J. Klippenstein, "Current status of transition-state theory," *J. Phys. Chem.* **100**, 12771–12800 (1996).
- ²¹K. J. Laidler and M. C. King, "Development of transition-state theory," *J. Phys. Chem.* **87**, 2657–2664 (1983).
- ²²N. E. Henriksen and F. Y. Hansen, *Theories of Molecular Reaction Dynamics: The Microscopic Foundation of Chemical Kinetics* (Oxford University Press, 2011).
- ²³C. F. Goldsmith and R. H. West, "Automatic generation of microkinetic mechanisms for heterogeneous catalysis," *J. Phys. Chem. C* **121**, 9970–9981 (2017).
- ²⁴L. Zhang, Y.-Q. Su, M.-W. Chang, I. A. W. Filot, and E. J. M. Hensen, "Linear activation energy-reaction energy relations for LaBO₃ (B = Mn, Fe, Co, Ni) supported single-atom platinum group metal catalysts for CO oxidation," *J. Phys. Chem. C* **123**, 31130–31141 (2019).
- ²⁵M. Besora and F. Maseras, "Microkinetic modeling in homogeneous catalysis," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**, e1372 (2018).
- ²⁶D. T. Gillespie, "Stochastic simulation of chemical kinetics," *Annu. Rev. Phys. Chem.* **58**, 35–55 (2007).
- ²⁷D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.* **81**, 2340–2361 (1977).
- ²⁸D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *J. Comput. Phys.* **22**, 403–434 (1976).
- ²⁹C. W. Gao, J. W. Allen, W. H. Green, and R. H. West, "Reaction Mechanism Generator: Automatic construction of chemical kinetic mechanisms," *Comput. Phys. Commun.* **203**, 212–225 (2016).
- ³⁰M. Keçeli, S. N. Elliott, Y.-P. Li, M. S. Johnson, C. Cavallotti, Y. Georgievskii, W. H. Green, M. Pelucchi, J. M. Wozniak, A. W. Jasper, and S. J. Klippenstein, "Automated computational thermochemistry for butane oxidation: A prelude to predictive automated combustion kinetics," *Proc. Combust. Inst.* **37**, 363–371 (2019).
- ³¹R. J. Gillis and W. H. Green, "Thermochemistry prediction and automatic reaction mechanism generation for oxygenated sulfur systems: A case study of dimethyl sulfide oxidation," *ChemSystemsChem* **2**, e1900051 (2020).
- ³²C. A. Grambow, A. Jamal, Y.-P. Li, W. H. Green, J. Zádor, and Y. V. Suleimanov, "Unimolecular reaction pathways of a γ -ketohydroperoxide from combined application of automated reaction discovery methods," *J. Am. Chem. Soc.* **140**, 1035–1048 (2018).
- ³³E. Martínez-Núñez, "An automated method to find transition states using chemical dynamics simulations," *J. Comput. Chem.* **36**, 222–234 (2015).
- ³⁴E. Martínez-Núñez, "An automated transition state search using classical trajectories initialized at multiple minima," *Phys. Chem. Chem. Phys.* **17**, 14912–14921 (2015).
- ³⁵J. A. Varela, S. A. Vázquez, and E. Martínez-Núñez, "An automated method to find reaction mechanisms and solve the kinetics in organometallic catalysis," *Chem. Sci.* **8**, 3843–3851 (2017).
- ³⁶A. Rodríguez, R. Rodríguez-Fernández, S. A. Vázquez, G. L. Barnes, J. J. P. Stewart, and E. Martínez-Núñez, "tssc2018: A code for automated discovery of chemical reaction mechanisms and solving the kinetics," *J. Comput. Chem.* **39**, 1922–1930 (2018).
- ³⁷S. Kopec, E. Martínez-Núñez, J. Soto, and D. Peláez, "vdw-TSSCDS—An automated and global procedure for the computation of stationary points on intermolecular potential energy surfaces," *Int. J. Quantum Chem.* **119**, e26008 (2019).
- ³⁸P. M. Zimmerman, "Automated discovery of chemically reasonable elementary reaction steps," *J. Comput. Chem.* **34**, 1385–1392 (2013).
- ³⁹P. Zimmerman, "Reliable transition state searches integrated with the growing string method," *J. Chem. Theory Comput.* **9**, 3043–3050 (2013).
- ⁴⁰J. P. Unsleber and M. Reiher, "The exploration of chemical reaction networks," *Annu. Rev. Phys. Chem.* **71**, 121 (2020).
- ⁴¹S. Maeda and K. Ohno, "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* **109**, 5742–5753 (2005).
- ⁴²K. Ohno and S. Maeda, "Automated exploration of reaction channels," *Phys. Scr.* **78**, 058122 (2008).
- ⁴³S. Maeda and K. Morokuma, "Finding reaction pathways of type A + B \rightarrow X: Toward systematic prediction of reaction mechanisms," *J. Chem. Theory Comput.* **7**, 2335–2345 (2011).
- ⁴⁴S. Habershon, "Sampling reactive pathways with random walks in chemical space: Applications to molecular dissociation and catalysis," *J. Chem. Phys.* **143**, 094106 (2015).
- ⁴⁵L.-P. Wang, A. Titov, R. McGibbon, F. Liu, V. S. Pande, and T. J. Martínez, "Discovering chemistry with an *ab initio* nanoreactor," *Nat. Chem.* **6**, 1044–1048 (2014).
- ⁴⁶B. Peters, A. Heyden, A. T. Bell, and A. Chakraborty, "A growing string method for determining transition states: Comparison to the nudged elastic band and string methods," *J. Chem. Phys.* **120**, 7877–7886 (2004).
- ⁴⁷X. Zhu, K. C. Thompson, and T. J. Martínez, "Geodesic interpolation for reaction pathways," *J. Chem. Phys.* **150**, 164103 (2019).
- ⁴⁸F. Perini, E. Galligani, and R. D. Reitz, "An analytical Jacobian approach to sparse reaction kinetics for computationally efficient combustion modeling with large reaction mechanisms," *Energy Fuels* **26**, 4804–4822 (2012).
- ⁴⁹C. K. Westbrook, W. J. Pitz, O. Herbinet, H. J. Curran, and E. J. Silke, "A comprehensive detailed chemical kinetic reaction mechanism for combustion of n-alkane hydrocarbons from n-octane to n-hexadecane," *Combust. Flame* **156**, 181–199 (2009).
- ⁵⁰J. A. Miller and C. F. Melius, "Kinetic and thermodynamic issues in the formation of aromatic compounds in flames of aliphatic fuels," *Combust. Flame* **91**, 21–39 (1992).
- ⁵¹J. Zeng, L. Cao, M. Xu, T. Zhu, and J. Z. H. Zhang, "Complex reaction processes in combustion unraveled by neural network-based molecular dynamics simulation," *Nat. Commun.* **11**, 5713 (2020).
- ⁵²S. Stocker, G. Csányi, K. Reuter, and J. T. Margraf, "Machine learning in chemical reaction space," *Nat. Commun.* **11**, 5505 (2020).
- ⁵³M. E. Jenkin, S. M. Saunders, and M. J. Pilling, "The tropospheric degradation of volatile organic compounds: A protocol for mechanism development," *Atmos. Environ.* **31**, 81–104 (1997).

- ⁵⁴K. Takahashi and I. Miyazato, "Rapid estimation of activation energy in heterogeneous catalytic reactions via machine learning," *J. Comput. Chem.* **39**, 2405–2408 (2018).
- ⁵⁵S. Choi, Y. Kim, J. W. Kim, Z. Kim, and W. Y. Kim, "Feasibility of activation energy prediction of gas-phase reactions by machine learning," *Chem. - Eur. J.* **24**, 12354–12358 (2018).
- ⁵⁶C. A. Grambow, L. Pattanaik, and W. H. Green, "Deep learning of activation energies," *J. Phys. Chem. Lett.* **11**, 2992–2997 (2020).
- ⁵⁷Kernel Methods for Activation Energy Prediction.
- ⁵⁸W. Jiang, X. Xing, X. Zhang, and M. Mi, "Prediction of combustion activation energy of NaOH/KOH catalyzed straw pyrolytic carbon based on machine learning," *Renewable Energy* **130**, 1216–1225 (2019).
- ⁵⁹J. Xu, X.-M. Cao, and P. Hu, "Improved prediction for the methane activation mechanism on rutile metal oxides by a machine learning model with geometrical descriptors," *J. Phys. Chem. C* **123**, 28802–28810 (2019).
- ⁶⁰D. C. Elton, Z. Boukouvalas, M. S. Butrico, M. D. Fuge, and P. W. Chung, "Applying machine learning techniques to predict the properties of energetic materials," *Sci. Rep.* **8**, 9059 (2018).
- ⁶¹C. A. Grambow, L. Pattanaik, and W. H. Green, "Reactants, products, and transition states of elementary chemical reactions based on quantum chemistry," *Sci. Data* **7**, 137 (2020).
- ⁶²J. N. Bronsted, "Acid and basic catalysis," *Chem. Rev.* **5**, 231–338 (1928).
- ⁶³M. G. Evans and M. Polanyi, "Inertia and driving force of chemical reactions," *Trans. Faraday Soc.* **34**, 11–24 (1938).
- ⁶⁴J. E. Sutton and D. G. Vlachos, "Effect of errors in linear scaling relations and Bronsted–Evans–Polanyi relations on activity and selectivity maps," *J. Catal.* **338**, 273–283 (2016).
- ⁶⁵J. K. Nørskov, T. Bligaard, B. Hvolbæk, F. Abild-Pedersen, I. Chorkendorff, and C. H. Christensen, "The nature of the active site in heterogeneous metal catalysis," *Chem. Soc. Rev.* **37**, 2163–2171 (2008).
- ⁶⁶J. A. Kammeraad, J. Goetz, E. A. Walker, A. Tewari, and P. M. Zimmerman, "What does the machine learn? Knowledge representations of chemical reactivity," *J. Chem. Inf. Model.* **60**, 1290–1301 (2020).
- ⁶⁷D. Rogers and M. Hahn, "Extended-connectivity fingerprints," *J. Chem. Inf. Model.* **50**, 742–754 (2010).
- ⁶⁸H. L. Morgan, "The generation of a unique machine description for chemical structures—A technique developed at chemical abstracts service," *J. Chem. Doc.* **5**, 107–113 (1965).
- ⁶⁹P. Schwaller, D. Probst, A. C. Vaucher, V. H. Nair, T. Laino, and J.-L. Reymond, "Data-driven chemical reaction classification, fingerprinting and clustering using attention-based neural networks," *Nat. Mach. Intell.* **3**(2), 144–152 (2021).
- ⁷⁰J. N. Wei, D. Duvenaud, and A. Aspuru-Guzik, "Neural networks for the prediction of organic chemistry reactions," *ACS Cent. Sci.* **2**, 725–732 (2016).
- ⁷¹N. Schneider, D. M. Lowe, R. A. Sayle, and G. A. Landrum, "Development of a novel fingerprint for chemical reactions and its application to large-scale reaction classification and similarity," *J. Chem. Inf. Model.* **55**, 39–53 (2015).
- ⁷²C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, Cambridge, MA, 2006).
- ⁷³V. S. Nguyen, H. L. Abbott, M. M. Dawley, T. M. Orlando, J. Leszczynski, and M. T. Nguyen, "Theoretical study of formamide decomposition pathways," *J. Phys. Chem. A* **115**, 841–851 (2011).
- ⁷⁴A. R. Singh, B. A. Rohr, J. A. Gauthier, and J. K. Nørskov, "Predicting chemical reaction barriers with a machine learning model," *Catal. Lett.* **149**, 2347–2354 (2019).
- ⁷⁵Q. Zhao, M. Avdeev, L. Chen, and S. Shi, "Machine learning prediction of activation energy in cubic Li-argyrodites with hierarchically encoding crystal structure-based (HECS) descriptors," *Sci. Bull.* **66**, 1401 (2021).
- ⁷⁶C. W. Coley, W. Jin, L. Rogers, T. F. Jamison, T. S. Jaakkola, W. H. Green, R. Barzilay, and K. F. Jensen, "A graph-convolutional neural network model for the prediction of chemical reactivity," *Chem. Sci.* **10**, 370–377 (2019).
- ⁷⁷G. Landrum, Rdkit: Open-source cheminformatics software, 2016.
- ⁷⁸F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.* **12**, 2825–2830 (2011); available at <http://jmlr.org/papers/v12/pedregosa11a.html>.
- ⁷⁹M. Stone, "Cross-validated choice and assessment of statistical predictions," *J. R. Stat. Soc. Ser. B* **36**, 111–147 (1974).
- ⁸⁰A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," *Phys. Rev. B* **87**, 184115 (2013).
- ⁸¹J. Zádor and H. N. Najm, "Kinbot 1.0: A code for automatic pes exploration," Technical Report, Sandia National Laboratories (SNL-CA), Livermore, CA, 2013.
- ⁸²A. Jalan, I. M. Alecu, R. Meana-Pañeda, J. Aguilera-Iparraguirre, K. R. Yang, S. S. Merchant, D. G. Truhlar, and W. H. Green, "New pathways for formation of acids and carbonyl products in low-temperature oxidation: The Korcek decomposition of γ -ketohydroperoxides," *J. Am. Chem. Soc.* **135**, 11100 (2013).
- ⁸³S. Ryu, Y. Kwon, and W. Y. Kim, "A Bayesian graph convolutional network for reliable prediction of molecular properties with uncertainty quantification," *Chem. Sci.* **10**, 8438–8446 (2019).
- ⁸⁴A. A. Peterson, R. Christensen, and A. Khorshidi, "Addressing uncertainty in atomistic machine learning," *Phys. Chem. Chem. Phys.* **19**, 10978–10985 (2017).
- ⁸⁵G. Scalia, C. A. Grambow, B. Pernici, Y.-P. Li, and W. H. Green, "Evaluating scalable uncertainty estimation methods for deep learning-based molecular property prediction," *J. Chem. Inf. Model.* **60**, 2697 (2020).