RESEARCH ARTICLE | MARCH 21 2025

Efficient exploration of reaction pathways using reaction databases and active learning

Domantas Kuryla ; Gábor Csányi ; Adri C. T. van Duin ; Angelos Michaelides

J. Chem. Phys. 162, 114122 (2025)
https://doi.org/10.1063/5.0235715

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Efficient exploration of reaction pathways using reaction databases and active learning

Cite as: J. Chem. Phys. 162, 114122 (2025); doi: 10.1063/5.0235715
Submitted: 29 August 2024 • Accepted: 24 February 2025 •
Published Online: 21 March 2025

Domantas Kuryla,[1] Gábor Csányi,[2,a)] Adri C. T. van Duin,[3] and Angelos Michaelides[1,a)]

AFFILIATIONS

[1] Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, United Kingdom
[2] Engineering Laboratory, University of Cambridge, Trumpington St and JJ Thomson Ave, Cambridge, United Kingdom
[3] Department of Mechanical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

[a)] **Authors to whom correspondence should be addressed:** gc121@cam.ac.uk and am452@cam.ac.uk

ABSTRACT

The fast and accurate simulation of chemical reactions is a major goal of computational chemistry. Recently, the pursuit of this goal has been aided by machine learning interatomic potentials (MLIPs), which provide energies and forces at quantum mechanical accuracy but at a fraction of the cost of the reference quantum mechanical calculations. Assembling the training set of relevant configurations is key to building the MLIP. Here, we demonstrate two approaches to training reactive MLIPs based on reaction pathway information. One approach exploits reaction datasets containing reactant, product, and transition state structures. Using an S_N2 reaction dataset, we accurately locate reaction pathways and transition state geometries of up to 170 unseen reactions. In another approach, which does not depend on data availability, we present an efficient active learning procedure that yields an accurate MLIP and converged minimum energy path given only the reaction end point structures, avoiding quantum mechanics driven reaction pathway search at any stage of training set construction. We demonstrate this procedure on an S_N2 reaction in the gas phase and with a small number of solvating water molecules, predicting reaction barriers within 20 meV of the reference quantum chemistry method. We then apply the active learning procedure on a more complex reaction involving a nucleophilic aromatic substitution and proton transfer, comparing the results against the reactive ReaxFF force field. Our active learning procedure, in addition to rapidly finding reaction paths for individual reactions, provides an approach to building large reaction path databases for training transferable reactive machine learning potentials.

I. INTRODUCTION

Chemical reactions are at the core of synthetic chemistry, catalysis, and biochemical processes. As a complement to experiments, computer simulation approaches have helped understand chemical reactions at the microscopic level. In the gas phase, this often involves the determination of the minimum energy path (MEP) and/or transition state (TS), from which reaction rates are subsequently predicted. In solution, solvent degrees of freedom need to be taken into consideration, and this can be done with enhanced sampling techniques such as metadynamics[1] and umbrella sampling.[2] However, accurate simulation results come at a large computational cost. Reaction pathway determination requires accurate force and total energy evaluation, typically obtained using quantum mechanics (QM) methods such as Density Functional Theory (DFT),[3,4] Hartree–Fock,[5] Møller–Plesset perturbation theory (MP2),[6] or coupled cluster.[7] For all QM methods, the computational cost scales steeply with system size, making solution phase reaction calculations particularly challenging. This, combined with the long simulation timescales required to sample solvent degrees of freedom, generally prohibits accurate large-scale simulations of reactions in solution. Therefore, there is a need for fast and accurate force and energy evaluation methods for complex chemical reactions.

Machine learning interatomic potentials (MLIPs) have enabled fast condensed phase simulations at QM accuracy. An MLIP is a machine learning architecture that maps a system's atomic positions to the system's potential energy, fitted to reproduce the potential energy surface (PES) of a chosen QM method.[8–11] Early MLIP

architectures include the Behler–Parrinello neural network potentials[12] and the Gaussian approximation potentials,[13] originally applied to study the properties of solids. Later, MLIPs were developed to study the properties of organic molecules in near-equilibrium geometries: to estimate molecular atomization energies[14] and relative stabilities of conformations and isomers.[15,16] MLIPs have enabled coupled cluster accuracy simulations of liquid water[17] and molecules[18,19] as well as DFT level simulations of biomolecules.[20] Only in recent years have approaches to modeling organic reactions with MLIPs been devised,[21–24] offering more flexible alternatives to empirical reactive force fields such as ReaxFF.[25] There are two main approaches to training reactive MLIPs: using preexisting datasets containing a wide range of reactions and building the training set for a specific reaction by sampling the relevant PES region. Considering the first approach, this has been done with a dataset comprising reactive configurations from nudged elastic band (NEB)[26] calculations: Schreiner *et al.* used the Transition1x dataset[27] to train a PaiNN[28] model and used the model to calculate NEB pathways of unseen reactions.[29] The Transition1x dataset was also used by Yuan *et al.* to train neural network potentials to predict Hessians for transition state optimization.[30] In addition, there are large reaction datasets that only comprise reactant, transition state, and product geometries and offer great potential in training reactive MLIPs. These include gas phase reaction datasets of 4466 E2 and S_N2 reactions;[31] 5000 cycloadditions;[32] 12 000 reactions including up to 7 C, N, and O atoms[33] (whose full reaction paths were calculated in the Transition1x dataset); 176 992 reactions including up to 10 C, N, and O atoms;[34] and monomolecular interconversions[35] involving molecules from the QM9 dataset.[36] Another approach to model reactions when the reactive geometries are not known in advance is to sample the relevant region of the PES using the active learning (AL) training technique.[37–40] AL starts with an initial training set, with the trained model used to sample new training points to iteratively refine the training set. The training set is enriched by sampling the PES using molecular dynamics,[41] downhill molecular dynamics initialized at the TS geometry,[42,43] displacing NEB images along normal modes,[44] or combinations of various techniques.[45,46] When training MLIPs for reactions in solution, enhanced sampling techniques such as metadynamics have been used to thoroughly sample the solvent and reacting species degrees of freedom.[47,48] For solution phase reactions, there is a strong motivation to build training sets efficiently due to expensive reference QM calculations. In pursuit of efficiency, an AL procedure has been developed in which new training data are selected using descriptor-based metrics, requiring fewer training points to achieve a stable and accurate MLIP compared to AL using energy error for training data selection.[49,50] A major goal in the development of reactive MLIPs is to provide models that accurately describe reactions in the condensed phase, such as reactions in solution. Procedures for fitting reactive MLIPs often take a hierarchical approach, first focusing on training data describing the reactive species and later the solvent–solute and solvent–solvent interactions. In this spirit, the present work is a stepping stone toward the overarching goal, with a focus on finding reaction pathways using MLIPs. In particular, it is desirable to develop ways to avoid computationally expensive QM calculations to sample the PES, which can instead be done using an MLIP, only using QM for single point calculations. In this work, we show that the reaction dataset and AL approaches can be efficient using the MACE[51] architecture. MACE has been shown to give stable potentials with relatively little data.[52,53] In the first approach, we show that pathways of unseen reactions can be found using models trained only using reactant, TS, and product geometries, avoiding computing NEB pathways purely using QM. For these models, we use the dataset of gas phase E2 and S_N2 reactions.[31] In the second approach, we obtain the NEB pathway and a reactive potential using AL, initializing the training set with structural information of reactants and products only. In the AL loop, the potential is used in a NEB calculation to include new configurations in the training set. This approach allows us to avoid the high computational cost of converging the NEB calculation with QM, whether the MEP is the end goal or is used as training data in more exhaustive configuration space sampling by AL, as done in Refs. 44 and 45.

This paper is structured as follows: In Sec. II, we explain the methods and data used in the work. In Sec. III A, we show the performance of MACE models trained on a reaction dataset in barrier prediction for reactions with unseen TS geometries and demonstrate using these models that TS geometries for unseen reactions can be obtained using NEB. In Sec. III B, we demonstrate our simple AL procedure with an S_N2 reaction in the gas phase and with four solvating water molecules, as well as a nucleophilic aromatic substitution reaction, where we can compare performance to the ReaxFF reactive empirical force field method using the Castro *et al.* CHNOS parameters[54] as well as our fine-tuned force field based on the Ashraf and van Duin C/H/O parameters.[55] In Sec. IV, we close with conclusions.

II. METHODS AND DATA

We now briefly describe the MACE machine learning architecture employed in this study, the reaction dataset, the quantum chemistry calculations, the NEB method, the ReaxFF simulations, and the active learning procedure.

A. MACE architecture

The MACE architecture[51] has been designed following developments in message passing for representations of atomic environments.[56,57] MACE maps the atomic positions and elements of a system to the total energy of the system. Decomposition of the total energy into atomic contributions enables linear scaling and the transferability of the model between systems with different numbers of atoms. The models used in this work are made up of two layers, each with a body order of four and a distance cutoff r_{max} of 5 Å, 128 embedding channels, and the highest order of the symmetric features $L_{max} = 1$. This way, the effective receptive field of each atom is up to 10 Å. In Sec. III B 3, models with 64 embedding channels were trained.

B. E2 and S_N2 reaction dataset

We trained MACE models on a database (QMrxn20) of E2 and S_N2 reactions[31] comprising reactant, product, and transition state structures of around 4500 reactions involving a base/nucleophile anion H^-, F^-, Cl^-, and Br^- attacking a derivatized haloethane (C_2H_5X with X = F, Cl, Br) species with up to two hydrogens on each carbon atom replaced with NO_2, CN, CH_3, and NH_2 groups. We chose this dataset because S_N2 and E2 reactions are commonly encountered in organic chemistry. In addition, we can easily study solvation effects since these reactions take place in the gas as well as

the solution phase. For building our training sets, we used all 4500 transition state structures and 4500 reactant complex structures.

C. Quantum chemistry calculations

In both parts of this work, we obtain the reference quantum chemistry energies and forces at the MP2/6-311G(d) level of theory, as implemented in ORCA 5.0.3.[58] This is the same level of theory as used in developing the original dataset in Ref. 31. In the MP2 calculations, all electrons were correlated (using the NoFrozencore option). The only noteworthy difference from Ref. 31 is that here we used the resolution of identity (RI) approximation[59] to MP2 with automatic generation of the auxiliary basis set. This approximation speeds up the QM calculation with little loss of accuracy. In particular, barrier heights with and without the RI approximation differed by 0.5 meV on average and by no more than 1 meV for 50 reactant and TS pairs selected from the dataset. We note that, while the choice of basis set can have an effect on predicted barrier heights, our main goal here is to train models that reproduce PESs of a chosen QM method rather than to use the most accurate available QM method. With this in mind, we used the same basis set as used in Ref. 31, while recognizing that higher level theories and better quality basis sets are available. In addition, when it came to the comparison of gas phase and microsolvated reactions in Secs. III B 1 and III B 2, we explored the sensitivity to basis set through calculations with a def2 TZVP basis set. With this higher quality basis set, we observed that the reaction energy changed, relative to the calculations with the 6-311G(d) basis set, by +0.51 and +0.07 eV for the gas phase and microsolvated reactions, respectively, while the forward barrier heights changed by +0.05 and −0.03 eV, respectively.

D. Nudged elastic band

We used the NEB method as implemented in the Atomic Simulation Environment.[60] When using MACE, we use NEB with no climbing image (unless stated otherwise), relying on increasing the number of images to obtain a better TS estimate if needed. The intermediate images are initialized by linearly interpolating between the initial and final images, while in Sec. III B 3, an image-dependent pair potential[61] was used to obtain a better initial path. The optimization is carried out using the FIRE optimization algorithm,[63] with a force stopping criterion of 0.05 eV/Å. The spring constant is scaled in proportion to the number of images, e.g., 0.2 eV/Å^2 for 20 images and 0.4 eV/Å^2 for 40 images. For NEB carried out purely using MP2, once the force on all images drops below 0.05 eV/Å, we switch to climbing image NEB[64] and further optimize with the same force stopping threshold.

E. ReaxFF reactive empirical force field

In this work, we used the ReaxFF reactive empirical force field using the Castro *et al.* coal combustion force field parameters.[54] We also used the Ashraf and van Duin C/H/O parameters.[55] ReaxFF NEB calculations were carried out using the LAMMPS software.[65] LAMMPS input details are given in Sec. S7 of the supplementary material.

F. Active learning procedure

The active learning procedure consists of training set initialization from reactant and product geometries and iterative improvement using NEB until a convergence criterion is met. We chose NEB for the sampling of reactive configurations for several reasons. First, it only requires the knowledge of reactant and product states. While the performance and, in some cases, the outcome of NEB depend on the initial guess of the pathway, there are unbiased ways to construct an initial pathway guess, including linear interpolation and image-dependent pair potential.[61] Second, NEB provides the entire path from the reactant state to the product state, making a broad range of reactive configurations available for selecting new training data. As illustrated in Fig. 1, the dataset is initialized by generating replicas of relaxed reactant and product structures and applying random

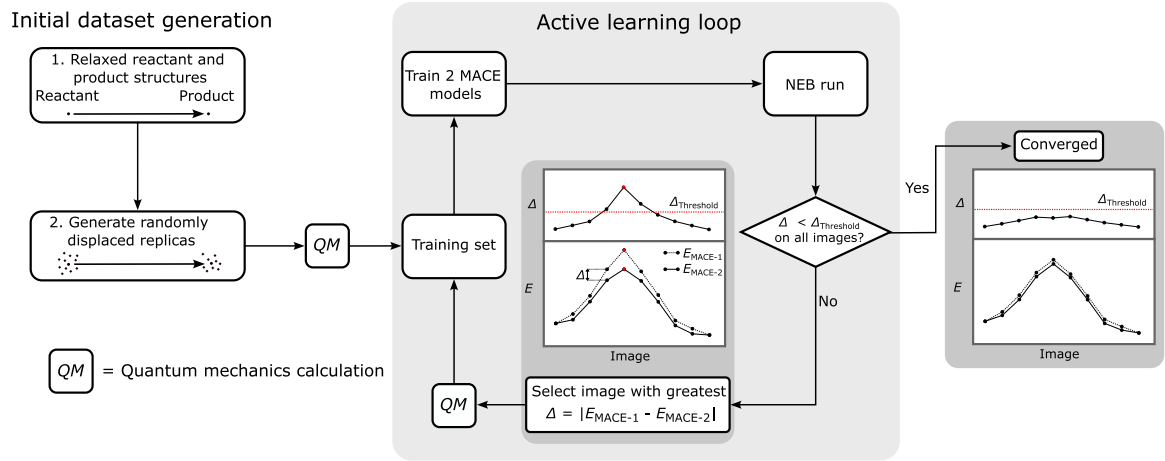

FIG. 1. Scheme of the active learning procedure. The training set is initialized by applying random displacements to the reactant and product structures. The active learning loop includes training the models, running a NEB calculation, selecting new configuration(s) based on model disagreement, calculating the reference energy and forces using the MP2 quantum mechanical approach, and updating the training set. If model disagreement is below a specified threshold on all NEB images, the active learning is considered converged.

Gaussian displacements to all atoms. A mean displacement of 0.02 Å was used. Tests with a 0.04 Å mean displacement were also carried out and are discussed in Sec. S6 of the supplementary material. We kept the displacements small to avoid overstretching or overcompressing covalent bonds, which would introduce outliers in the small training set. Since the same average displacement magnitude is applied to all atoms, regardless of their role in the reaction, we expect these displacement magnitudes to be reasonable for many reactions. For each end point structure (reactant and product), we generate $N_{\text{train}} + N_{\text{valid}}$ displaced structures, which are then split among the training and validation sets. This way, the initial training set is composed of equal parts of reactant and product displaced structures. The validation set is also composed of equal parts of displaced reactant and product structures. The purpose of the validation set is that the resulting model saved is from the epoch with the lowest loss on the validation set. In this work, we initialized the training set and the validation set with several tens of configurations each (e.g., 40, 60), consisting of equal parts of reactant and product configurations with random atomic displacements.

In the active learning loop, two models are trained using the same training set but initialized with different random weights. We also did tests with a committee of four models, where the disagreement is the standard deviation of energy predictions in the committee. We run a NEB calculation using the average forces of the models. The NEB is optimized for up to 250 steps. If NEB has not converged in 250 steps and model disagreements on images are below the threshold, there is an option to continue running NEB until an image with a large disagreement is encountered. This option was used in Sec. III B 3, in addition to the improved tangent method[62] and climbing image NEB to refine the transition state region. Selection of new training data uses final NEB images with model disagreement exceeding the threshold. QM energies and forces are evaluated in the order of decreasing model disagreement until a specified number (e.g., 1, 5) of labeled structures are evaluated. If reference forces exceed a threshold of 10 eV/Å on any atom, that image is discarded, and the next image with the highest disagreement is considered. With the training set updated, the models are retrained. The AL loop is considered converged if the model can converge the NEB and the model disagreement is below the threshold on all final images. Detailed input arguments are given in Sec. S5 of the supplementary material.

III. RESULTS

A. Unseen barrier and TS geometry prediction with MACE trained on a reaction dataset

Reaction datasets can provide diverse data for training reactive MLIPs. The minimal composition of many reaction datasets, including only the reactant, transition state, and product structures per reaction, means that it is more challenging to use them to train stable MLIPs. Here, we demonstrate that, using the MACE architecture, we can efficiently train reactive MLIPs that can accurately predict unseen reaction barriers and be used to find reaction pathways.

1. Iterative training to predict unseen TS energies and reaction barriers

We trained a MACE model on a dataset containing reactants and transition states of various E2 and S_N2 reactions in the QMrxn20 dataset.[31] Reaction barriers in the dataset range between −2 and 4 eV. For some reactions, the corresponding reactant structures are highly strained as a result of poor optimization. For those reactions, the energy difference between the transition state and the reactant state is negative. With the goal of minimizing the number of transition states used for the training, we trained an initial model using 4500 reactant structures provided in the dataset. The training set was then iteratively expanded by including small numbers of transition state structures. In each iteration, 1% of transition states with the highest energy errors are transferred to the training set.

The performance of the model in predicting barrier height (the difference between the TS energy and the reactant energy given their structures) at different stages of iterative training is shown in Fig. 2. A model trained with no TS structures in the training set shows a root mean squared error (RMSE) of more than 300 meV. This is a large error, but the prediction is sufficiently accurate that a correlation between the MP2 and MACE barrier heights is seen. After more than 5% of transition states were included in the training set, the RMSE in unseen reaction barrier heights dropped below 43 meV or 1 kcal/mol. This value is widely quoted as the "chemical accuracy" limit since the errors below 43 meV do not affect rate and equilibrium constants significantly. Throughout the iterative training, errors go down across the entire test set, as shown in Fig. S1 in the supplementary material. We continued the iterative training until

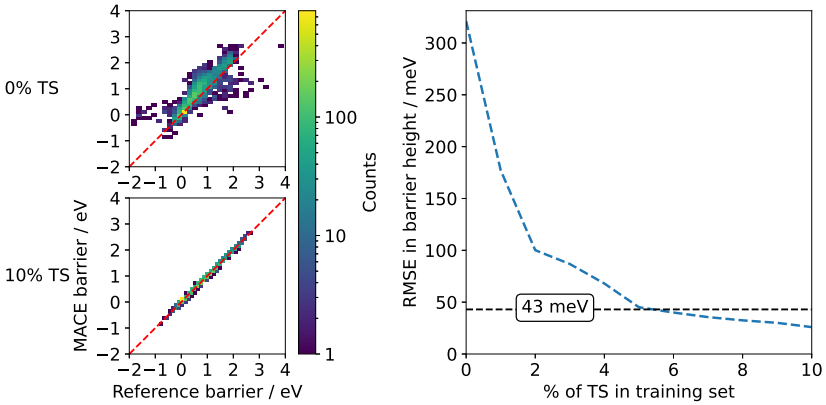

FIG. 2. Performance of E2 and S_N2 barrier prediction with models trained on only reactant structures (0% transition states) and on increasing fractions of transition states included in the training set. A "chemical accuracy" of 43 meV in RMSE on unseen barriers was reached after iteratively including more than 5% of the transition state structures in the training set. The energies are reported in eV, with 1 eV = 23.06 kcal/mol.

10% of the TS were in the training set, where the RMSE on reaction barriers with unseen TS was 26 meV. Therefore, overall we find that a model trained only on the reactant structures showed qualitative accuracy on barrier heights, while a model trained on all reactant and a small fraction of TS structures showed small errors on barrier heights in reactions where the TS is not in the training set. While this approach is generalizable to other reaction datasets, the amount of required training data can vary. In particular, for more diverse reaction datasets, we expect that a larger fraction of transition states would need to be used for training to accurately predict remaining barriers.

2. TS geometry prediction with NEB

While high accuracy on unseen TS structures is a sign that predictions can be made with chemical accuracy using the MLIP, in the general case of modeling reactions, the TS geometry is not known for a given reactant–product pair without performing expensive reaction pathway search calculations with QM methods. Therefore, potentials are most useful in practice if they can be used to locate reaction paths.

To test the performance of our models in finding reaction paths, we used the QMrxn20 dataset to generate a separate dataset of reactant, transition state, and product structures for S_N2 reactions involving halogen nucleophiles and leaving groups (reactions involving the hydride ion as the nucleophile were excluded). Geometry optimization starting from the TS structures was performed to ensure conformational consistency between the reactant, product, and TS structures. The reactant and product structures were optimized until the forces on all atoms were below 0.1 eV/Å. As before, the MP2/6-311G(d) level of theory was used in the optimization. Where the optimization yielded incorrect structures (e.g., deprotonation by the S_N2 nucleophile), the reactions were discarded, leaving a total of 1778 S_N2 reactions, each with their respective reactant, product, and transition state structures. We trained a new MACE model using training, validation, and test splits of 1422, 178, and 178 reactions, respectively. We also altered the training set by applying random atomic displacements (standard deviations of 0.02 and 0.05 Å) to the training set structures and trained respective models. Before carrying out NEB calculations, we checked the accuracy of the models. The RMSE in barrier heights, obtained by evaluating the MACE model on known test set structures, for the same test reactions were 35, 16, and 31 meV for the models trained with 0, 0.02, and 0.05 Å random displacements in the training set, respectively. The errors show that the models trained are transferable to the unseen reactions.

With these three models, we carried out NEB calculations for 178 unseen reactions in the test set. The highest MACE energy image was taken as the predicted TS. To assess the quality of the predicted TS geometries, we compute errors as the differences in MP2 energy between the predicted and the true TS geometries. These errors are summarized in Fig. 3 for the model trained with no random atomic displacements (MACE-sn2) and the model trained with atomic displacements with σ = 0.02 Å (MACE-sn2-rattled-A). The results from all three models, including MACE-sn2-rattled-B, trained with atomic displacements of σ = 0.05 Å, are given in detail in Table S1 of the supplementary material. While the majority of the predicted TS fall close in energy to the true TS, some NEB calculations resulted in large errors. This is also reflected in inset (a). For the MACE-sn2 model, the RMSE on TS energies was, excluding 15 unconverged NEB paths, 0.946 eV, with 65% of TS predictions having absolute errors below 0.05 eV. The MACE-sn2-rattled-A model performed better on the same reactions, leading to an overall RMSE on TS energies of 0.195 eV, excluding 3 unconverged NEB paths, with 93% of TS estimates within 0.05 eV of the true TS. Thus, adding random displacements to the training set atoms yielded a more stable potential and a larger fraction of converged reaction

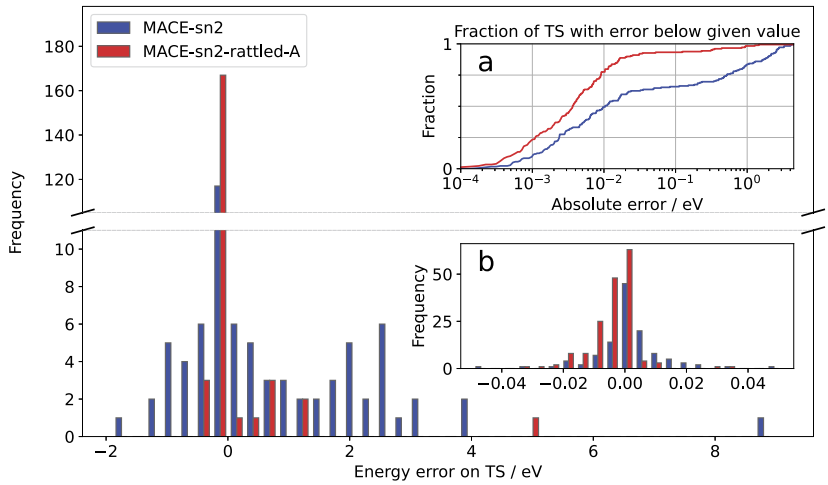

FIG. 3. Predicted TS MP2 energy error as the difference between the energy of the TS structure obtained from NEB (as the highest energy image) and the energy of the true TS structure from the 178 reactions in the test set. Positive (negative) error values mean that the predicted TS has a higher (lower) energy than the true TS. Inset (a) shows the number of predicted TS with absolute energy error below a given value. Inset (b) is an expansion of the region in TS energy error close to 0, between −0.05 and 0.05 eV, where the majority of predicted TS lie. Applying random atomic displacements (rattling) to the training set structures improves the fraction of accurate TS guesses and reduces the fraction of unphysical pathways. The energies are reported in eV, with 1 eV = 23.06 kcal/mol.

pathways. A possible cause for this result is that random atomic displacements provide information about the shape of the PES further away from the local minima and the TS. We illustrate this in Fig. S2 of the supplementary material, where we calculated the PES, with the reaction coordinates constrained and other coordinates relaxed, for one unseen reaction with the three models. MACE-sn2 reproduced the shape of the PES around the reaction pathway but showed incorrectly low energies for configurations where the bonds are too stretched or too compressed. Thus, the model has stability issues. MACE-sn2-rattled-A and MACE-sn2-rattled-B accurately predicted the energy in regions around the MEP and the steep climb in energy further away from the MEP.

Our work provides an efficient approach for reaction pathway prediction. Without computing full reaction pathways for training data, we were able to train potentials that can accurately find the reaction pathway in the majority of unseen reactions. We used 5334 configurations in the training set and used the resulting model for 178 unseen reactions, each taking a few thousand single point gradients to obtain the converged reaction path. However, the dataset used here comprises only one reaction class, which makes it easier to achieve transferability to unseen reactions. A similar approach in training transferable reactive MLIPs was shown by Schreiner *et al.*[29] They trained a PaiNN model using the Transition1x dataset containing reaction pathways of 10 000 reactions involving systems of up to 7 non-hydrogen atoms (C, N, O). The model was used as a surrogate potential for determining pathways of unseen reactions, with an RMSE of 0.24 eV on predicted barrier heights. Due to the different datasets and MLIP architectures used here and in Ref. 29, we cannot make direct comparisons of performance. Thus, it is an open question how our approach of only training with equilibrium and TS geometries performs in more chemically diverse reaction datasets.

B. Reactive MLIP via active learning

Reaction databases, such as the one discussed in Sec. III A 2, cover a small portion of chemical reaction space. In particular, they generally do not include configurations of solvated reacting species. Therefore, it is desirable to have an approach that does not depend on a large pre-computed database. Ideally, it would also not rely on an expensive QM-driven initialization of the training set. In this section, we demonstrate a simple active learning procedure, described in Sec. II F, that achieves these two goals. We show that this procedure allows us to locate the MEP with chemical accuracy, requiring very few QM single point evaluations (<100) for a gas phase S_N2 reaction as part of the QMrxn20 dataset. Next, we show that AL can be used in a similar manner to find the reaction path in a simplified model of a solvated reaction for the same reaction with four solvating water molecules. Finally, we showcase our AL method on a nucleophilic aromatic substitution reaction, comparing it with results using the ReaxFF potential. In this section, we focus on efficiently training potentials that accurately describe the minimum energy paths of reactions. Such potentials are mostly applicable to gas phase reactions. However, we included a reaction with a small number of solvent molecules as a test to extend the method to solvated reactions. This test is an initial step toward our overarching goal to develop potentials suitable for reactions in condensed phase. This includes reactions in solution, where solvent degrees of freedom need to be treated explicitly, and is part of our ongoing work.

1. Ethyl chloride and fluoride gas phase S_N2 reaction

We applied the AL procedure to a simple, prototypical gas phase S_N2 reaction between ethyl chloride and a fluoride anion. It took 9 AL iterations to obtain a model that converges the MEP

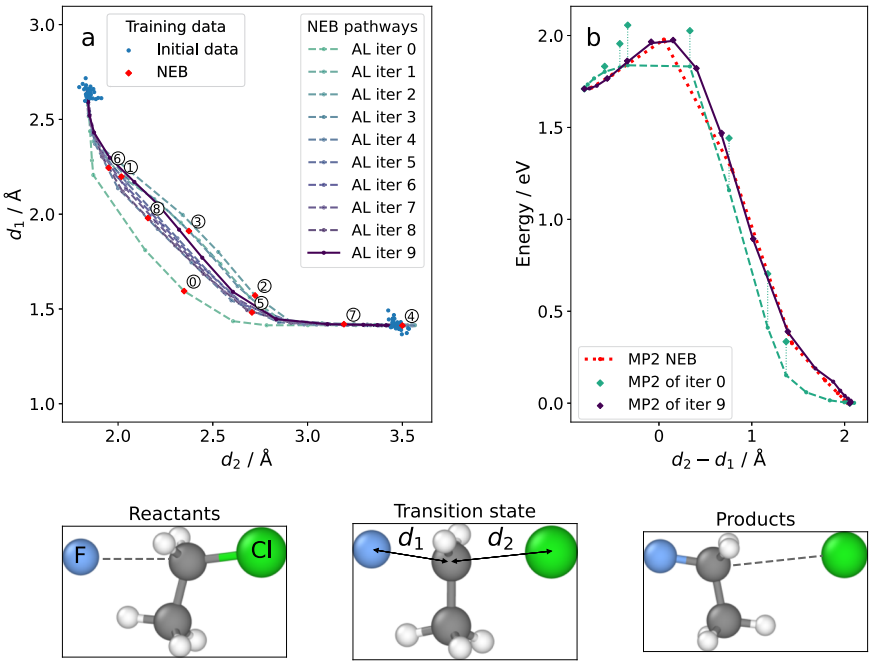

FIG. 4. Active learning on gas phase S_N2 reaction between ethyl chloride and a fluoride ion. The bottom panels show the structures of reactants, products, and transition state, indicating the reaction coordinates d_1 and d_2. Panel (a) shows, projected on the reaction coordinates, the training data (initial and obtained in active learning). Data points selected from NEB pathways are indicated with their respective iteration number. Panel (b) shows MACE energy profiles of selected NEB pathways and their corresponding MP2 energies indicated in diamond markers of the same color, along with the pathway obtained using climbing image NEB with MP2. The energies are reported in eV, with 1 eV = 23.06 kcal/mol.

with the disagreement between the MACE models not exceeding the threshold of 1.0 meV per atom on all images, or 9.0 meV overall. We use reaction coordinates d_1 and d_2, respectively the distances from the fluorine and chlorine atoms to the reacting carbon atom, to illustrate the training results in Fig. 4. From this, we learned that the initial model, trained using near-equilibrium geometries, is stable enough to provide a reasonable guess of the MEP. Throughout the AL iterations, a different MEP guess is obtained, which allows us to sample different regions of the reaction coordinate space, as shown in panel (a) of Fig. 4. Once the AL was converged, the MEP was found with energy errors $|E_{\mathrm{MACE}} - E_{\mathrm{MP2}}|$ of no more than 11 meV on any image, as shown in panel (b). From this MEP, the reaction barrier of 0.261 eV was predicted with MACE. The model error means that the MACE MEP differs from the MEP on the ground truth PES. To quantify this difference, we calculated the MEP purely with MP2 NEB to find a barrier of 0.262 eV.

We note that here, the goal of the AL procedure is to obtain the MEP accurately and predict reaction barriers. For comparison, other AL procedures for generating reactive potentials acquire training data beyond the MEP to predict free energy surfaces[44] and reaction product ratios.[43] In the case of potentials trained in our AL procedure, the high accuracy is limited to the reaction pathways. In AL, the model disagreement on and around the MEP is low but rises further away from the MEP. In Fig. S3 of the supplementary material, we show the PES calculated for the ethyl chloride–fluoride system, highlighting the region of the PES where the energy error relative to MP2 is below 40 meV, or roughly 1 kcal/mol. It can be seen that the low error region is in the vicinity of the minimum energy pathway.

With our AL procedure, we can replace the computationally expensive MP2 calculations in the MEP determination with the cheap MACE potential, requiring only the geometry information of the reactants and products to start with. For comparison, converging the NEB with MP2 required 62 NEB iterations, totaling 558 MP2 evaluations for nine intermediate NEB images, while 69 evaluations were required to construct the training set, in addition to 60 configurations for the validation set.

2. Solvation effects on S_N2 reaction barrier found with active learning

The simulation of reactions in solution is a highly sought-after application of MLIPs. While finding reaction barriers in solution phase reactions involves sampling over many degrees of freedom of the reacting species and the solvent molecules, we can apply the NEB active learning procedure to simplified solvation model systems with a limited number of solvent molecules.

We constructed a simplified model of solvation of the S_N2 reaction discussed in Sec. III B 1. The microsolvated reactant and product structures were constructed by placing four water molecules: two next to the chlorine and two next to the fluorine, with one hydrogen atom of each water molecule facing the nearby halogen atom. We then relaxed these structures at the MP2/6-311G(d) level of theory, stopping the optimization when the forces on all atoms are below 0.05 eV/Å.

The results of the training are shown in Fig. 5. A total of 15 additional training structures obtained from the NEB procedure were required to converge the active learning using a model disagreement threshold of 1.0 meV per atom (21.0 meV for the full system). In the first few AL iterations, the NEB pathway visited high energy regions, either with both d_1 and d_2 values being small or both being large. After a few high energy configurations were included, the NEB was steered toward a realistic reaction pathway involving concerted C–F bond formation and C–Cl bond breaking. The final iteration MEP has errors of up to 15 meV, lying close to the true MEP on the reference PES. In total, 75 MP2 single-point evaluations were

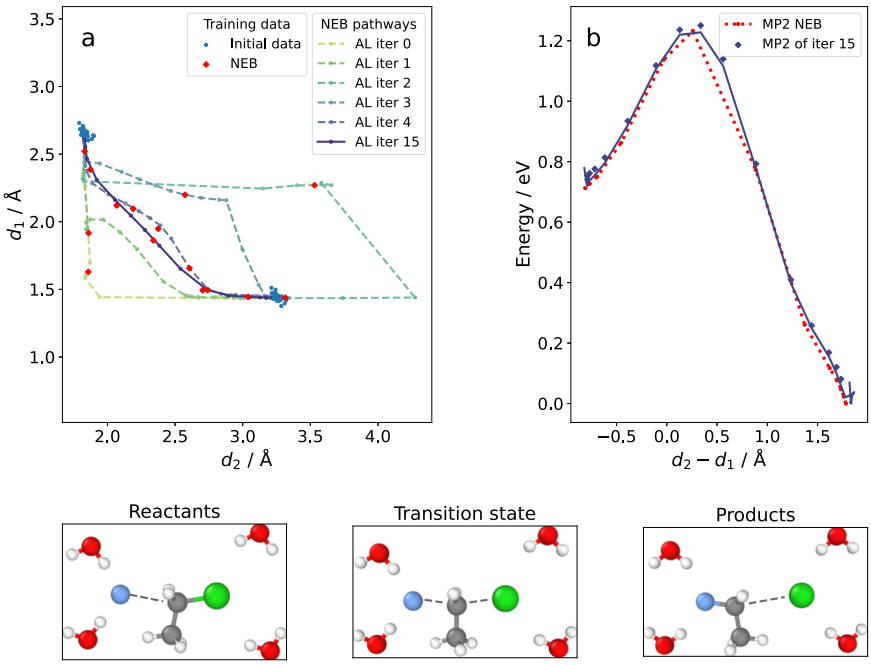

FIG. 5. Active learning on microsolvated S_N2 reaction between a fluoride ion and ethyl chloride. Throughout the active learning, the predicted pathway covers a broad range in the reaction coordinate space, allowing us to collect different configurations in the training set, as shown in panel (a). Panel (b) shows the MACE energy profile of the final iteration NEB pathway and its corresponding MP2 energies indicated in diamond markers of the same color, along with the pathway obtained using climbing image NEB with MP2. The energies are reported in eV, with 1 eV = 23.06 kcal/mol.

carried out to construct the training set and 60 for the validation set. For comparison, converging a NEB with 13 intermediate images at the MP2 level of theory required 78 optimizer steps, a total of 1014 single-point evaluations.

While using MACE in NEB is orders of magnitude faster than using QM, model training time is an important consideration as it can also be a computationally time-consuming step. We find that we can optimize the training time by increasing the number of new data points selected in each AL iteration. Including more new configurations in the training set in each iteration can make the AL converge in fewer iterations. The effect of more than one update configuration per iteration on training time and final training set size is shown in Sec. S6 of the supplementary material. In particular, we find that by accepting up to five images per iteration, the AL converged in six iterations and 28 added NEB geometries in less than 2 h on a single GPU core, while acceptance of one geometry per iteration required 15 iterations and added NEB geometries and 5 h to converge AL. However, for large systems, where the QM calculation time dominates, it is beneficial to only add one new geometry to the training set per AL iteration. Thus, there is an efficiency trade-off in the AL process that depends on the relative costs of the training and the QM components.

Before moving to Sec. III B 3, we briefly discuss interesting differences observed between the gas phase and microsolvated reaction. The final MEP of the microsolvated reaction shows a higher energy barrier for the fluoride to react with the ethyl chloride (0.50 eV) compared to the unsolvated reaction (0.26 eV). In addition, the energy difference between the reactants and products is reduced from 1.71 eV in the unsolvated system to 0.75 eV in the microsolvated system. These results are similar to those reported in the QM/MM study of the S_N2 reaction between methyl chloride and a fluoride ion in gas and in aqueous phase.[66] There, the effect of solvation was shown to be the increase in the reaction barrier and reduction in exothermicity. The results are consistent with the fluoride ion being a stronger hydrogen bond acceptor than the chloride ion. While the reactants are stabilized with respect to the products, reducing the reactant–product energy difference, the stabilization of the reactants with respect to the transition state increases the barrier height of the forward reaction. Thus, the AL procedure provides an efficient way to evaluate solvation effects by using simplified model systems.

3. Nucleophilic aromatic substitution involving proton transfer

Finally, we briefly showcase the AL approach developed here on a more complex reaction. In particular, we investigate a nucleophilic aromatic substitution reaction between 2-methylthiopyridine and methanol.

We also use this reaction as an opportunity to compare the AL approach to empirical reactive force fields. In particular, we compare to the well-known ReaxFF[25] family of potentials. ReaxFF uses a bond-order formalism to allow for changes in atomic connectivity throughout the simulation. ReaxFF parameters are typically trained against quantum mechanical data describing reaction energies and reaction barriers. This training happens prior to molecular dynamics simulations; no on-the-fly quantum mechanics simulations are performed. This enables ReaxFF simulations for large numbers of atoms (≫10 000 atoms) and timescales (>tens of nanoseconds). ReaxFF potentials may not accurately capture reactions that they were not explicitly trained for. Here, we compare the results of NEB carried out using ReaxFF parameters which were trained for the combustion of coal[54] and not explicitly for aromatic substitution reactions,

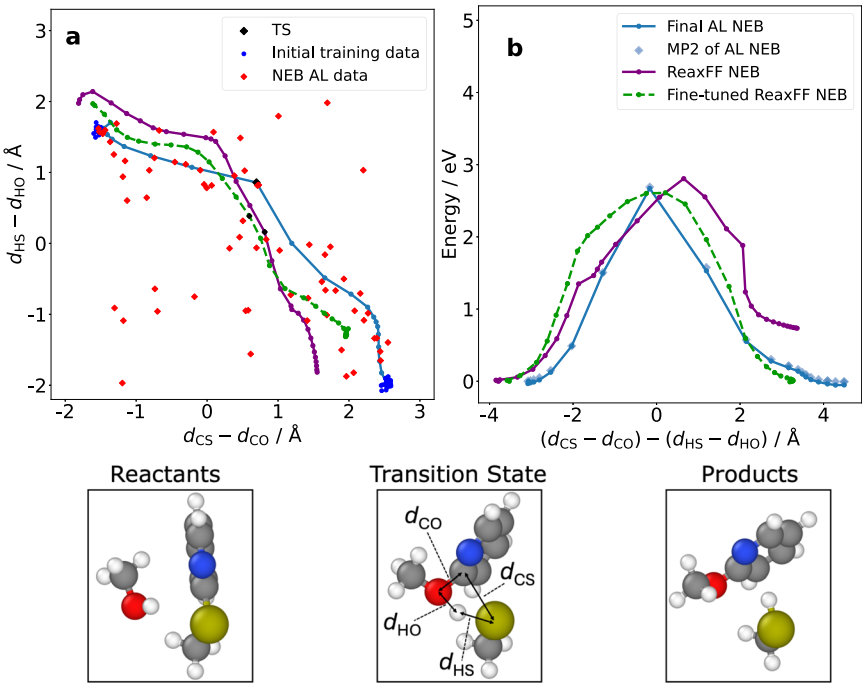

FIG. 6. NEB pathways obtained for an aromatic substitution reaction using AL as well as a ReaxFF force field including the coal combustion parameters from Ref. 54 and a fine-tuned force field based on the C/H/O parameters in Ref. 55. Panel (a) displays the AL training points and NEB pathways along the reaction coordinates. Panel (b) shows the energy profiles along a combined reaction coordinate. The images at the bottom of the figure show the structures of the reactant, transition state, and product. The energies are reported in eV, with 1 eV = 23.06 kcal/mol.

and the NEB AL procedure on an aromatic substitution between methanol and 2-(methylthio)pyridine. As illustrated in Fig. 6, the reaction involves two key processes: displacement of the methylthio group sulfur by methanol oxygen and proton transfer from the oxygen atom to the sulfur atom. These processes are reflected as two reaction coordinates. NEB with MACE trained on MP2 predicts these two processes to happen simultaneously. ReaxFF predicts an earlier nucleophilic attack and a later proton transfer. The NEB AL, initialized with 40 training points, required an additional 67 configurations from NEB to converge. For comparison, converging a NEB with eight intermediate images, the MP2 level of theory required 120 optimizer steps, a total of 960 single-point evaluations. While ReaxFF qualitatively agrees with MACE trained on MP2 on the forward barrier (2.81 and 2.67 eV for ReaxFF and MACE, respectively), it predicts the product energy 0.79 eV higher than MACE does, hence it disagrees on the reverse barrier (2.07 and 2.72 eV). The active learning, being more computationally demanding than the ReaxFF NEB (hours using AL vs seconds using ReaxFF), provides a more accurate prediction of the reaction pathway. Since the ReaxFF parameters used here were not explicitly trained for this reaction, we also explored fine-tuning a ReaxFF model to our reaction. Specifically, we used a CHNOS force field based on the Ashraf 2017 C/H/O parameters, which were not changed during the retraining process. We retrained the parameters that control interactions between C, H, S atoms using a coal combustion dataset associated with Ref. 54. We then added structures of our reaction from the MACE NEB path to that training set and retrained the same parameters. We recomputed the NEB path using the fine-tuned ReaxFF model.

The resulting reaction path is shown in panel (a) and the energy profile in panel (b) of Fig. 6, matching the MACE barrier height by 50 meV and reaction energy by 3 meV. The geometry of the fine-tuned ReaxFF pathway is closer to that of the MACE pathway. However, a more complete training dataset would be needed for the ReaxFF model to reproduce the MACE PES.

IV. CONCLUSIONS

Finding reaction pathways purely with a quantum chemistry level of theory is computationally expensive. With MLIPs, we can circumvent this high computational cost given relevant training data. Even with minimal reaction datasets containing reactants, transition states, and products, we can train MLIPs to accurately locate reaction pathways for many unseen reactions, while previous approaches in the literature required training data with entire reaction pathways. With active learning, we built MLIPs for finding MEPs with chemical accuracy efficiently, only using QM for selected new training points.

Our long-term goal is to train transferable reactive potentials that are applicable to condensed phase reactions. In an initial step toward this goal, we focused on efficiently training reactive potentials for finding the minimum energy path. We believe that the approaches demonstrated to efficiently find reaction pathways will be useful in building training sets for general reactive organic MLIPs. A possible avenue for general reactive MLIPs involves the use of foundation models fine-tuned using reaction data. The foundation models are trained on condensed phase non-reactive structures, such as the model by Batatia *et al.*,[53] providing qualitative accuracy for solvated systems. With this strategy toward general reactive MLIPs, the ability to quickly obtain reaction pathways for numerous reactions in the gas phase is highly useful, as we can fine-tune the foundation model to get accurate reactive behavior while retaining the accuracy of the interactions between solvent molecules and the reactive species.

SUPPLEMENTARY MATERIAL

See the supplementary material for additional data supporting discussion and statements made in the main text.

ACKNOWLEDGMENTS

This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital Grant No. EP/T022159/1), and DiRAC funding from the Science and Technology Facilities Council (www.dirac.ac.uk).DK acknowledges the support from the EPSRC Center for Doctoral Training in Chemical Synthesis with Grant Reference No. EP/S024220/1. A.M. acknowledges support from the European Union under the "n-AQUA" European Research Council project (Grant No. 101071937). A.C.T.v.D. acknowledges support from award DMR # 2306042 from the U.S. National Science Foundation.

AUTHOR DECLARATIONS

Conflict of Interest

GC has an equity interest in Symmetric Group LLP that licenses force fields commercially and also in Ångstrom AI.

Author Contributions

Domantas Kuryla: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Visualization (equal); Writing – original draft (equal). **Gábor Csányi**: Conceptualization (equal); Formal analysis (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal). **Adri C. T. van Duin**: Conceptualization (equal); Writing – review & editing (equal). **Angelos Michaelides**: Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The input files associated with this study and all analysis can be found on GitHub at https://github.com/water-ice-group/organic_neb_mace. We also include any datasets built in this work and used to train models.

REFERENCES

[1] A. Laio and M. Parrinello, "Escaping free-energy minima," Proc. Natl. Acad. Sci. U. S. A. **99**(20), 12562–12566 (2002).
[2] G. M. Torrie and J. P. Valleau, "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling," J. Comput. Phys. **23**(2), 187–199 (1977).

[3]P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," Phys. Rev. **136**(3B), B864–B871 (1964).
[4]W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," Phys. Rev. **140**, A1133–A1138 (1965).
[5]D. Rayner Hartree and W. Hartree, "Self-consistent field, with exchange, for beryllium," Proc. R. Soc. London, Ser. A **150**(869), 9–33 (1935).
[6]M. Head-Gordon, J. A. Pople, and M. J. Frisch, "MP2 energy evaluation by direct methods," Chem. Phys. Lett. **153**(6), 503–506 (1988).
[7]Č. Jiří, "On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods," J. Chem. Phys. **45**(11), 4256–4266 (1966).
[8]S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, "Comparing molecules and solids across structural and alchemical space," Phys. Chem. Chem. Phys. **18**(20), 13754–13769 (2016).
[9]A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi, and M. Ceriotti, "Machine learning unifies the modeling of materials and molecules," Sci. Adv. **3**(12), e1701816 (2017).
[10]V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, "Gaussian process regression for materials and molecules," Chem. Rev. **121**(16), 10073–10141 (2021).
[11]F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, and M. Ceriotti, "Physics-inspired structural representations for molecules and materials," Chem. Rev. **121**(16), 9759–9815 (2021).
[12]J. Behler and M. Parrinello, "Generalized neural-network representation of high-dimensional potential-energy surfaces," Phys. Rev. Lett. **98**(14), 146401 (2007).
[13]A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons," Phys. Rev. Lett. **104**(13), 136403 (2010).
[14]M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, "Fast and accurate modeling of molecular atomization energies with machine learning," Phys. Rev. Lett. **108**, 058301 (2012).
[15]J. S. Smith, O. Isayev, and A. E. Roitberg, "ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost," Chem. Sci. **8**(4), 3192–3203 (2017).
[16]D. Péter Kovács, C. van der Oord, J. Kucera, A. E. A. Allen, D. J. Cole, C. Ortner, and G. Csányi, "Linear atomic cluster expansion force fields for organic molecules: Beyond RMSE," J. Chem. Theory Comput. **17**(12), 7696–7711 (2021).
[17]J. Daru, H. Forbert, J. Behler, and D. Marx, "Coupled cluster molecular dynamics of condensed phase systems enabled by machine learning potentials: Liquid water benchmark," Phys. Rev. Lett. **129**(22), 226001 (2022).
[18]J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev, and A. E. Roitberg, "Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning," Nat. Commun. **10**(1), 2903 (2019).
[19]P. O. Dral, A. Owens, A. Dral, and G. Csányi, "Hierarchical machine learning of potential energy surfaces," J. Chem. Phys. **152**(20), 204110 (2020).
[20]O. T. Unke, M. Stöhr, S. Ganscha, T. Unterthiner, H. Maennel, S. Kashubin, D. Ahlin, M. Gastegger, L. Medrano Sandonas, J. T. Berryman, A. Tkatchenko, and K.-R. Müller, "Biomolecular dynamics with machine-learned quantum-mechanical force fields trained on diverse chemical fragments," Sci. Adv. **10**(14), eadn4397 (2024).
[21]Y. Yang, S. Zhang, K. D. Ranasinghe, O. Isayev, and A. E. Roitberg, "Machine learning of reactive potentials," Annu. Rev. Phys. Chem. **75**, 371–395 (2024).
[22]S. Stocker, G. Csányi, K. Reuter, and J. T. Margraf, "Machine learning in chemical reaction space," Nat. Commun. **11**(1), 5505 (2020).
[23]S. Stocker, H. Jung, G. Csányi, C. Franklin Goldsmith, K. Reuter, and J. T. Margraf, "Estimating free energy barriers for heterogeneous catalytic reactions with machine learning potentials and umbrella integration," J. Chem. Theory Comput. **19**(19), 6796–6804 (2023).
[24]E. Gelžinytė; M. Öeren, M. D. Segall, and G. Csányi, "Transferable machine learning interatomic potential for bond dissociation energy prediction of drug-like molecules," J. Chem. Theory Comput. **20**(1), 164–177 (2024).
[25]T. P. Senftle, S. Hong, Md. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, and A. C. T. van Duin, "The ReaxFF reactive force-field: development, applications and future directions," npj Comput. Mater. **2**(1), 15011 (2016).
[26]G. Mills and H. Jónsson, "Quantum and thermal effects in H_2 dissociative adsorption: Evaluation of free energy barriers in multidimensional quantum systems," Phys. Rev. Lett. **72**(7), 1124–1127 (1994).
[27]M. Schreiner, A. Bhowmik, T. Vegge, J. Busk, and O. Winther, "Transition1x—A dataset for building generalizable reactive machine learning potentials," Sci. Data **9**(1), 779 (2022).
[28]K. Schütt, O. Unke, and M. Gastegger, "Equivariant message passing for the prediction of tensorial properties and molecular spectra," Proc. Mach. Learn. Res. **139**, 9377–9388 (2021).
[29]M. Schreiner, A. Bhowmik, T. Vegge, P. B. Jørgensen, and O. Winther, "NeuralNEB—Neural networks can find reaction paths fast," Mach. Learn.: Sci. Technol. **3**(4), 045022 (2022).
[30]E. C.-Y. Yuan, A. Kumar, X. Guan, E. D. Hermes, A. S. Rosen, J. Zádor, T. Head-Gordon, and S. M. Blau, "Deep learning of ab initio Hessians for transition state optimization," Nat. Commun. **15**, 8865 (2024).
[31]G. F. von Rudorff, S. N. Heinen, M. Bragato, and O. A. von Lilienfeld, "Thousands of reactants and transition states for competing E2 and S_N2 reactions," Mach. Learn.: Sci. Technol. **1**(4), 045026 (2020).
[32]T. Stuyver, K. Jorner, and C. W. Coley, "Reaction profiles for quantum chemistry-computed [3 + 2] cycloaddition reactions," Sci. Data **10**, 66 (2023).
[33]C. A. Grambow, L. Pattanaik, and W. H. Green, "Reactants, products, and transition states of elementary chemical reactions based on quantum chemistry," Sci. Data **7**(1), 137 (2020).
[34]Q. Zhao, S. M. Vaddadi, M. Woulfe, L. A. Ogunfowora, S. S. Garimella, O. Isayev, and B. M. Savoie, "Comprehensive exploration of graphically defined reaction spaces," Sci. Data **10**, 145–4463 (2023).
[35]S. Nandi, T. Vegge, and A. Bhowmik, "MultiXC-QM9: Large dataset of molecular and reaction energies from multi-level quantum chemical methods," Sci. Data **10**(1), 783 (2023).
[36]R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, "Quantum chemistry structures and properties of 134 kilo molecules," Sci. Data **1**(1), 140022 (2014).
[37]J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev, and A. E. Roitberg, "Less is more: Sampling chemical space with active learning," J. Chem. Phys. **148**(24), 241733 (2018).
[38]C. Schran, K. Brezina, and O. Marsalek, "Committee neural network potentials control generalization errors and enable active learning," J. Chem. Phys. **153**(10), 104105 (2020).
[39]N. Bernstein, G. Csányi, and V. L. Deringer, "De novo exploration and self-guided learning of potential-energy surfaces," npj Comput. Mater. **5**(1), 99 (2019).
[40]C. van der Oord, M. Sachs, D. P. Kovács, C. Ortner, and G. Csányi, "Hyperactive learning for data-driven interatomic potentials," npj Comput. Mater. **9**(1), 168 (2023).
[41]S. Zhang, M. Z. Makoś, R. B. Jadrich, E. Kraka, K. Barros, B. T. Nebgen, S. Tretiak, O. Isayev, N. Lubbers, R. A. Messerly, and J. S. Smith, "Exploring the frontiers of condensed-phase chemistry with a general reactive machine learning potential," Nat. Chem. **16**(5), 727–734 (2024).
[42]T. A. Young, T. Johnston-Wood, V. L. Deringer, and F. Duarte, "A transferable active-learning strategy for reactive molecular force fields," Chem. Sci. **12**(32), 10944–10955 (2021).
[43]T. A. Young, T. Johnston-Wood, H. Zhang, and F. Duarte, "Reaction dynamics of Diels–Alder reactions from machine learned potentials," Phys. Chem. Chem. Phys. **24**(35), 20820–20827 (2022).
[44]K. Brezina, H. Beck, and O. Marsalek, "Reducing the cost of neural network potential generation for reactive molecular systems," J. Chem. Theory Comput. **19**(19), 6589–6604 (2023).
[45]S. J. Ang, W. Wang, D. Schwalbe-Koda, S. Axelrod, and R. Gómez-Bombarelli, "Active learning accelerates ab initio molecular dynamics on reactive energy surfaces," Chem **7**, 738–751 (2021).
[46]L. L. Schaaf, E. Fako, S. De, A. Schäfer, and G. Csányi, "Accurate energy barriers for catalytic reaction pathways: An automatic training protocol for machine learning force fields," npj Comput. Mater. **9**(1), 180 (2023).
[47]M. Yang, L. Bonati, D. Polino, and M. Parrinello, "Using metadynamics to build neural network potentials for reactive events: The case of urea decomposition

in water," Catal. Today **387**, 143–149 (2022), part of Special Issue: 100 Years of CASALE SA: A Scientific Perspective on Catalytic Processes.
[48]V. Vitartas, H. Zhang, V. Juraskova, T. Johnston-Wood, and F. Duarte, "Active learning meets metadynamics: Automated workflow for reactive machine learning potentials," chemRxiv:10.26434.
[49]H. Zhang, V. Juraskova, and F. Duarte, "Modelling chemical processes in explicit solvents with machine learning potentials," Nat. Commun. **15**(1), 6114 (2024).
[50]V. Jurásková, G. Tusha, H. Zhang, L. V. Schäfer, and F. Duarte, "Modelling ligand exchange in metal complexes with machine learning potentials," Faraday Discuss. **256**, 156–176 (2025).
[51]I. Batatia, D. P. Kovács, G. Simm, C. Ortner, and G. Csányi, "Mace: Higher order equivariant message passing neural networks for fast and accurate force fields," in *Advances in Neural Information Processing Systems*, edited by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Curran Associates, Inc., 2022), Vol. 35, pp. 11423–11436.
[52]D. P. Kovács, I. Batatia, E. S. Arany, and G. Csányi, "Evaluation of the MACE force field architecture: From medicinal chemistry to materials science," J. Chem. Phys. **159**(4), 044118 (2023).
[53]I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, M. Avaylon, W. J. Baldwin, F. Berger, N. Bernstein, A. Bhowmik, S. M. Blau, V. Cărare, J. P. Darby, S. De, F. Della Pia, V. L. Deringer, R. Elijošius, Z. El-Machachi, F. Falcioni, E. Fako, A. C. Ferrari, A. Genreith-Schriever, J. George, R. E. A. Goodall, C. P. Grey, P. Grigorev, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. Holm, J. Jaafar, S. Hofmann, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, J. R. Kermode, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, J. T. Margraf, I.-B. Magdău, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. L. Schaaf, C. Schran, B. X. Shi, E. Sivonxay, T. K. Stenczel, V. Svahn, C. Sutton, T. D. Swinburne, J. Tilly, C. van der Oord, E. Varga-Umbrich, T. Vegge, M. Vondrák, Y. Wang, W. C. Witt, F. Zills, and G. Csányi, "A foundation model for atomistic materials chemistry," arXiv: 2401.00096 [physics.chem-ph] (2024).
[54]F. Castro-Marcano, A. M. Kamat, M. F. Russo, C. T. van Duin, and J. P. Mathews, "Combustion of an Illinois No. 6 coal char simulated using an atomistic char representation and the ReaxFF reactive force field," Combust. Flame. **159**(3), 1272–1285 (2012).
[55]C. Ashraf and A. van Duin, "Extension of the ReaxFF combustion force field towards syngas combustion and initial oxidation kinetics," J. Phys. Chem. A **121**(5), 1051 (2017).
[56]J. Nigam, S. Pozdnyakov, G. Fraux, and M. Ceriotti, "Unified theory of atom-centered representations and message-passing machine-learning schemes," J. Chem. Phys. **156**(20), 204115 (2022).
[57]A. Bochkarev, Y. Lysogorskiy, C. Ortner, G. Csányi, and R. Drautz, "Multilayer atomic cluster expansion for semilocal interactions," Phys. Rev. Res. **4**(4), L042019 (2022).
[58]F. Neese, F. Wennmohs, U. Becker, and C. Riplinger, "The ORCA quantum chemistry program package," J. Chem. Phys. **152**(22), 224108 (2020).
[59]X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler, "Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, RPA, MP2 and *GW* with numeric atom-centered orbital basis functions," New J. Phys. **14**(5), 053020 (2012).
[60]A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. . B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, "The atomic simulation environment—A Python library for working with atoms," J. Phys.: Condens. Matter **29**(27), 273002 (2017).
[61]S. Smidstrup, A. Pedersen, K. Stokbro, and H. Jonsson, "Improved initial guess for minimum energy path calculations," J. Chem. Phys. **140**, 214106 (2014).
[62]G. Henkelman and H. Jónsson, "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points," J. Chem. Phys. **113**, 9978–9985 (2000).
[63]E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch, "Structural relaxation made simple," Phys. Rev. Lett. **97**(17), 170201 (2006).
[64]G. Henkelman, B. P. Uberuaga, and H. Jónsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," J. Chem. Phys. **113**(22), 9901–9904 (2000).
[65]A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, "LAMMPS—A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales," Comput. Phys. Commun. **271**, 108171 (2022).
[66]J. Zhang, Y. Xu, J. Chen, and D. Wang, "A multilayered-representation, quantum mechanical/molecular mechanics study of the $CH_3Cl + F^-$ reaction in aqueous solution: The reaction mechanism, solvent effects and potential of mean force," Phys. Chem. Chem. Phys. **16**(16), 7611–7617 (2014).