Check for updates

Reaction dynamics of Diels–Alder reactions from machine learned potentials†

Tom A. Young, [ID] * Tristan Johnston-Wood, [ID] Hanwen Zhang [ID] and
Fernanda Duarte [ID] *

Recent advances in the development of reactive machine-learned potentials (MLPs) promise to transform reaction modelling. However, such methods have remained computationally expensive and limited to experts. Here, we employ different MLP methods (ACE, NequIP, GAP), combined with automated fitting and active learning, to study the reaction dynamics of representative Diels–Alder reactions. We demonstrate that the ACE and NequIP MLPs can consistently achieve chemical accuracy (± 1 kcal mol^{-1}) to the ground-truth surface with only a few hundred reference calculations. These strategies are shown to enable routine *ab initio*-quality classical and quantum dynamics, and obtain dynamical quantities such as product ratios and free energies from non-static methods. For ambimodal reactions, product distributions were found to be strongly dependent on the QM method and less so on the type of dynamics propagated.

1 Introduction

Simulating chemical reactions is essential to developing a fundamental understanding of their mechanism and predicting experimental outcomes.[1] Machine learned potentials (MLPs) offer an enticing approach to such simulations, enabling the efficient mapping between nuclear configurations and energies ($R \mapsto E$). Moreover, in contrast to classical force fields, they offer flexibility and systematic improvability.[2,3] In the limit of correct forces and converged sampling, such simulations should afford access to the accurate estimation of rate and equilibrium constants. However, despite the development of Gaussian Approximation Potentials (GAPs)[4,5] and high dimensional neural network potentials (NNPs)[6] more than ten years ago, MLPs are still yet to find routine use for chemical reaction simulation.[7] This slow uptake is likely due to the computational and time investment required to train reactive potentials for new system, with only a handful of examples reported to date; these include S$_N$2,[8] combustion,[9] pericyclic,[10] decomposition,[11] dissociation,[12] proton transfer[13] and photochemical reactions.[14,15]

Iterative training is essential to construct robust MLPs and can be summarised as follows: (i) developing a training set, (ii) performing the regression, (iii) repeating the process until the desired accuracy is obtained. Automated approaches to construct large training sets have been previously reported;[16–18] however, their application remain limited to small systems. Moreover, the need to perform several thousands of energy and gradient evaluations on a reference potential energy surface (PES) makes it difficult to go beyond density functional theory (DFT) and employ accurate but computationally expensive wavefunction (WF)-based methods.[19] Exceptions are rare and limited to systems with $\lesssim 10$ atoms.[17,20,21] Finally, achieving the desired accuracy when training reactive MLPs is challenging because the energy scale required for accurate dynamics is much larger than for non-reactive MLPs (where sampling around the minima is often sufficient). As such, more training data is required to reach the same accuracy. This is particularly relevant when sampling areas around the transition states (TSs), which often require WF-based methods, such as coupled-cluster (CC).[22] Here, we show that recently developed MLP methods[23,24] can be used to generate accurate potentials for modestly sized reactions (≈ 50 atoms) in an automated fashion and outline some of the enabled insights.

2 Results and discussion

2.1 MLP comparison for a Diels–Alder reaction

We previously introduced a strategy to generate reactive GAPs in an autonomous manner, requiring only hundreds to a few thousand energy and gradient evaluations.[17] Here, we extend this methodology to explore more complex processes and compare it to recently developed MLP approaches (Fig. 1). As a starting point, we considered Diels–Alder (DA) reactions

Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, UK.
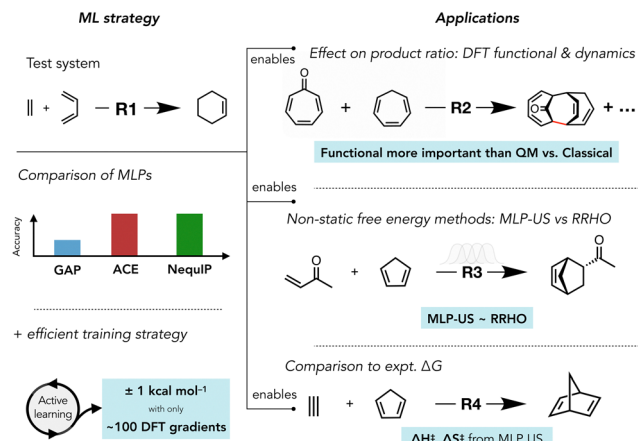*E-mail: tom.a.young@ucl.ac.uk, fernanda.duartegonzalez@chem.ox.ac.uk*

Fig. 1 Diels–Alder reactions studied in this work, summarising the different aspects explored with the trained MLPs.

because of the available theoretical and experimental data,[25,26] and their prominence in chemical and biochemical contexts.[27–29]

Initial efforts to apply our original active learning (AL) strategy using GAP regression to model the reaction between ethene + butadiene, **R1**, proved promising, producing qualitatively reasonable reactive molecular dynamics (MD). However, the quality of these GAPs were not within 1 kcal mol^{-1} accuracy of the QM reference method, which is required for accurate rate estimation or dynamic studies (Fig. S1a, MAD = 5.3 kcal mol^{-1}, ESI†). We hypothesised that this was due to the exothermicity of this reaction, which requires the MLP to be accurate in a 60

kcal mol^{-1} energy window. Indeed, using the same strategy and hyperparameters to model a less exothermic H-abstraction reaction ($H_3C^\bullet + C_3H_8 \rightarrow CH_4 + {}^\bullet CH(CH_3)_2$) resulted in a more accurate potential (Fig. S1b, MAD = 2.1 kcal mol^{-1}, ESI†). A modest improvement was obtained upon hyperparameter optimisation, at moderate computational cost (≈ 500 configurations required for **R1**). Specifically, reducing the regularisation, increasing the quality of the radial basis, and doubling the number of atomic environments considered in the training all improved the GAP (ESI,† Section S2). However, attempts to improve GAPs by training a component-wise potential separated over covalent bonds were unsuccessful (ESI,† Section S3.3).

We then turned our attention to the recently developed Atomic Cluster Expansion (ACE[23]) and Neural Equivariant Interatomic Potentials (NequIP[24]) MLP methods and evaluated their performance within the same training strategy for **R1** (Fig. 2). While rather different in philosophy, both ACE and NequIP outperform GAPs and provide MLPs of similar accuracy (Fig. 2a and Fig. S24, ESI†). Here, accuracy is based on deviations between QM ground-truth energies (true) and predicted energies over independent DFT-MD trajectories propagated from the TS to the reactant and product states. Previously, we have shown that a prospective validation strategy in the configuration space accessible to that MLP is essential to characterising 'good' MLPs.[17] These potentials are stable by construction as we use an AL strategy (Fig. 2, bottom left), where we define a stable potential where 10×1 ps trajectories can be propagated without encountering a configuration that deviates >2.3 kcal mol^{-1} (0.1 eV) from the true energy.
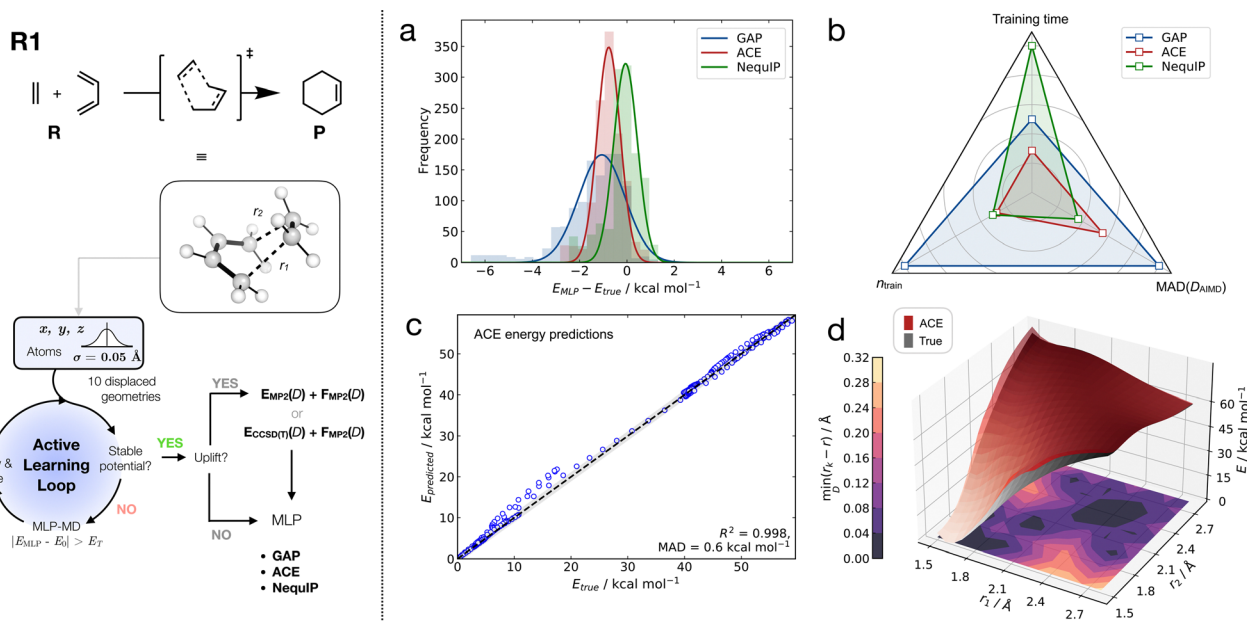

Fig. 2 MLP methods trained on the [4+2] DA cycloaddition between ethene and butadiene in the gas phase. (a) Signed errors on total energies over three trajectories from reactants to products (*ab initio* AIMD ground truth, 300 K, PBE0/def2-SVP). (b) Relative performance between MLP methods on total time, data efficiency (number of training configurations, n_{train}), and accuracy. Data normalised to unity (raw data in ESI,† Table S4). (c) Parity plot between MP2/def2-TZVP total energies and ACE predictions on MP2 AIMD trajectories from the TS. ACE trained on DFT-selected configurations with energies and forces re-evaluated at MP2. (d) Comparison of the predicted (red) and true PBE0/def2-SVP relaxed 2D potential energy surface surrounding the TS. Contour plot represents the 'closeness' of the training data to a point in the surface, with the darker regions being the closest.

This journal is © the Owner Societies 2022

Phys. Chem. Chem. Phys., 2022, **24**, 20820–20827 | **20821**

When considering the amount of data required to train an accurate MLP for **R1**, GAPs both with and without hyperparameter optimisation required ≈ 500 configurations. However, this is surpassed by both ACE and NequIP potentials, where $n_{train} \approx 100$ (Fig. 2b). Comparing training time across these methods is less straightforward due to their different architecture and implementation. In all cases they are found to be efficient and therefore suitable for rapidly developing bespoke MLPs. Among them, NequIP had the greatest training time (half a day using 10 CPU cores + 1 GPU), while GAP and ACE required just 5 ± 2 h of total training time on 10 CPU cores. It is important to note that this difference in training time reduces with the system size, as reference energy and force evaluations dominate the computational time for equally data efficient MLPs.

Tangent to our goal of developing accurate MLPs for DA reactions, we found that GAP regularisation could be harnessed to reduce the computational cost of developing CCSD(T)-quality potentials (ESI,† Section S4). For simple molecules, MP2 forces are accurate to within 1.1 kcal mol^{-1} Å^{-1} (0.05 eV Å^{-1}) of their CCSD(T) counterparts (Fig. S16, ESI†), thus within the 'expected error' of the GAP. This removes the requirement for numerical CCSD(T) gradients. This strategy, combined with CUR[30] selection, afford a 100-fold reduction in the number of required CCSD(T) energy calculations, thus greatly reducing the time required to train a CCSD(T)-quality potential (Table S3, ESI†). Furthermore, re-evaluating energies and forces from AL configurations with a different reference method (*i.e.* 'uplifting') reduces the computational cost associated with training WF-quality MLPs. Such strategies were also applied to other electronic structure methods. For example, uplifting PBE0/DZ configurations to MP2/TZ affords an ACE MLP within ± 1 kcal mol^{-1} of MP2-MD-sampled configurations (Fig. 2c, DZ = double-ζ basis set, TZ = triple-ζ). These strategies enable us to attain a higher level of accuracy in the MLP at a reduced computational cost.

Comparison of the ACE-, GAP- and NequIP-predicted two-dimensional PES over the two forming C–C bonds in **R1**, where all other degrees of freedom are free to relax, reveals that all three potentials are accurate and smooth in the extrapolation regime (Fig. S26, ESI†). While all three potentials are suitable for training this reaction, ACE offers the best overall balance in terms of data efficiency, training time and accuracy of the final potential (Fig. 2b). Moreover, even when the closest configuration in the training data is 0.3 Å away in the forming bonds, the error is only a couple of kcal mol^{-1} when AL is initiated at the TS ($r_1 = r_2 = 2.30$ Å), suggesting that ACE is accurate in the extrapolation regime (Fig. 2d). Based on these results, in the following sections we focus our discussions on ACE due to its overall efficiency and accuracy.

2.2 Product ratios and time gaps

To demonstrate the applicability of the ACE framework to more complex reactions, we modelled the reaction between tropone and cyclohepatriene, **R2**, and monitored its product distribution. This reaction has been previously studied by Houk and co-workers using quasi-classical reaction dynamic simulations.[31] The reaction has been found to involve an ambimodal TS that leads to three distinct products: one [8+2] and two [6+4] cycloadducts (Fig. 3, top).
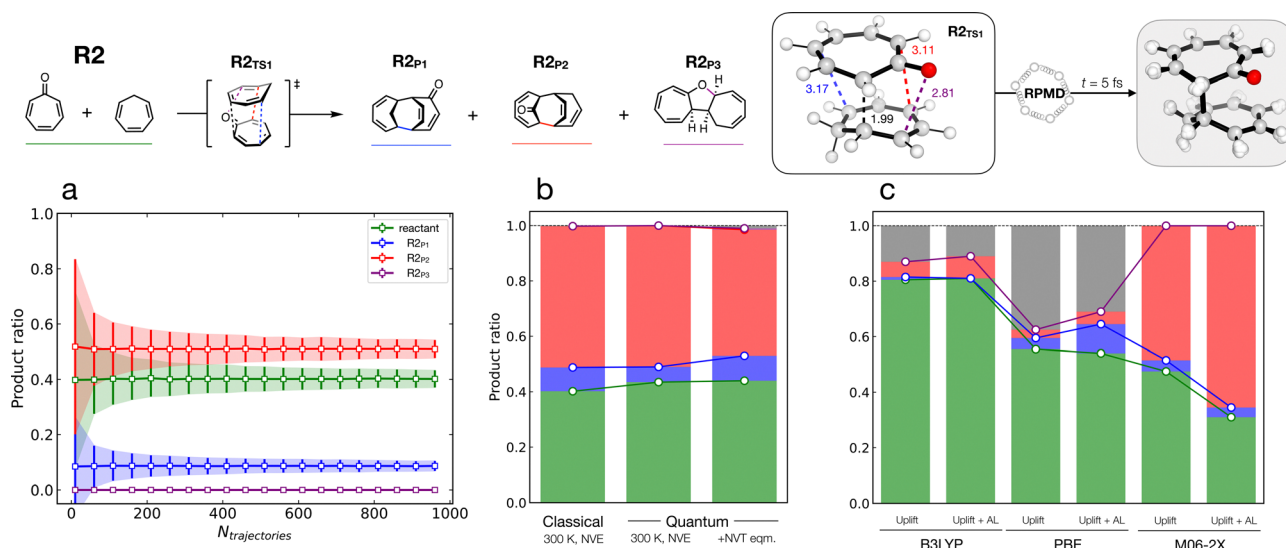

Fig. 3 Product distributions for the reaction between tropone and cycloheptatriene by ACE MD simulations initiated from a single TS (left). The ACE MLP was trained using a standard AL loop (diff, E_T = 2.3 kcal mol^{-1} = 0.1 eV, 500 K dynamics) from two TSs, such that the training data included all three products. (a) Convergence of the product ratio (*e.g.* $N_{\rightarrow R2_{P1}}/N_{total}$) with the number of trajectories propagated. Error is quoted as 2σ, obtained from bootstrapping with resampling (1000 iterations) on the same dataset. MD propagated from the TS in the *NVE* ensemble using initial Boltzmann-sampled velocities for 300 K. (b) Dependence of the product ratio on the type of propagated dynamics (N_{total} = 200), including classical and ring polymer MD (RPMD) from 1 ps simulations. (c) Product ratios from 200 classical *NVE* trajectories propagated from TS1 using ACE MLPs trained at B3LYP, PBE and M06-2X levels of theory (def2-SVP). The grey color refers to a no defined state being formed in 1 ps. Uplift corresponds to single point energy and force evaluations on ACE AL configurations (PBE0/def2-SVP) and then retraining the ACE potential. Uplift + AL uses PBE0/def2-SVP configurations reevaluated plus 5 iterations of AL at 500 K from the corresponding TS. See Methods sections S4 and S8 (ESI†) for more details.

20822 | *Phys. Chem. Chem. Phys.*, 2022, **24**, 20820–20827

This journal is © the Owner Societies 2022

Initial training of an ACE potential from the TS ($R2_{TS1}$) generated ≈ 450 configurations using standard AL with sampling in the reactant and product regions of two of the products ($R2_{P1}$ and $R2_{P2}$). However, even when propagating MLP-MD trajectories at 500 K, the Cope product ($R2_{P3}$) was not observed in the training data. Only when the TS that leads directly to $R2_{P3}$ was included as an AL starting configuration was the training data sufficiently complete. This second TS was 4.4 kcal mol^{-1} higher in energy than the TS that leads to $R2_{P1}$ and $R2_{P2}$.[31] This result highlights the importance of considering relevant points in the PES that may not be obtained using automated sampling methods when training MLPs. For this reaction, ACE training required ≈ 2 days of AL.

Having successfully trained this ACE MLP for **R2**, we propagated dynamics to compute product distribution. Using 100 trajectories, each of them taking only minutes, we were able to obtain converged results (Fig. 3a). This efficiency also allowed us to propagate ring polymer MD (RPMD)[32] using the developed ACE potential. As expected, given the nature of the bonds being formed during this reaction, the product ratios were similar to those obtained using classical MD.

We also evaluated the effect of temperature by equilibrating the system at the TS using constrained *NVT* dynamics (1 eV Å^{-2}, ≈ 0.5 kcal mol^{-1} additional ZPE), prior to propagating *NVE* downhill dynamics. Both direct *NVE* downhill and equilibrated dynamics provide qualitatively similar trends. However, a more detailed analysis shows that the latter affords almost double the proportion of $R2_{P1}$ and the formation of $R2_{P3}$ as 1% of the total product distribution, which would otherwise not be observed (right most distribution in Fig. 3b). This demonstrates the importance of propagating equilibrated quantum dynamics to obtain converged distributions.[33] Training an MLP uniquely allows for the 100-fold more expensive simulations to be performed, with the effect of ZPE and tunnelling accounted for.

Finally, we evaluated the effect of QM methodology on product distributions by performing MLP uplifts to different levels of theory. Uplift corresponds to performing single point energy and force evaluations on ACE AL configurations (PBE0/ def2-SVP) at different levels of theory and then retraining the ACE potential to propagate classical *NVE* trajectories from the respective TSs. For the product distributions in Fig. 3c, MLPs can be obtained at the desired level of theory in 100s CPUh by applying uplifts, compared with $>200\,000$ CPUh without uplifts. It was found that some levels of theory require more AL iterations than others, suggesting that the overlap of the configuration space within the training set may not be optimal. For example, while the product distributions are unchanged for B3LYP on additional AL, M06-2X distributions continue changing upon further AL iterations, affording $\approx 10\%$ more product $R2_{P2}$ after 5 iterations. Product distributions vary considerably between QM methodology, with the major state varying from the reactant to $R2_{P2}$, and the proportion of $R2_{P1}$ varying from 1% to 10% (Fig. 3c). This evidences the large influence that the level of theory has on the propagated dynamics and consequently on the product distribution. The differences are
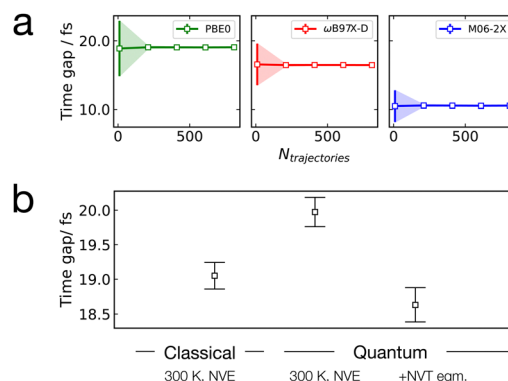

Fig. 4 Computed time gap for the reaction between methyl-vinyl ketone (MVK) and cyclopentadiene (CPD), **R3**. (a) Convergence of the time gap with the number of trajectories, propagated using ACE potentials trained at different levels of theory. Dynamics propagated classically in the *NVE* ensemble initiated from TSs of the corresponding functionals, with initial velocities obtained from a Boltzmann distribution at 300 K. (b) Dependence of time gaps on type of dynamics. MLP MD, referred to as Classical, were propagated for 200 trajectories initialised from the TS. MLP RPMD, referred to as Quantum, were performed in the *NVE* ensemble for 200 fs with and without *NVT* RPMD equilibration. Error are quoted as 2σ, obtained from bootstrapping with resampling on the same dataset.

significantly more important than those observed when comparing classical MD and RPMD results.

To further investigate the influence of the level of theory and the type of dynamics, the time gap, defined as the time difference between the formation of two C–C σ-bonds,[34] is evaluated for the reaction of methyl-vinyl ketone (MVK) + cyclopentadiene (CPD), **R3**. ACE MLPs trained at different levels of theory led to slightly different time gaps. For example, both PBE0 and ωB97X-D functionals generated average time gaps of ≈ 20 fs, while for M06-2X it was ≈ 10 fs (Fig. 4a). The latter is consistent with the DFT-MD value of 11 ± 6 fs in the literature,[35] indicating the reaction is dynamically concerted, as concluded by Houk and co-workers.[34,36] Furthermore, as expected, quantum effects on this reaction are minimal, (Fig. 4b), since the differences between classical MD and RPMD are within 2 fs, while the differences due to the level of theory vary by over 5 fs. Traditionally, such estimates require several hundred quasi-classical trajectories, which are computationally demanding. However, by employing our strategy, we can compute the time gap from 1000 trajectories in fewer than 6 hours.

2.3 Free energies

We further demonstrate the applicability of the ACE MLPs by employing them to compute free energy contributions for the reaction between MVK and CPD, **R3**. These contributions are often calculated within the rigid-rotor-harmonic-oscillator (RRHO) approximation. However, this method is known to be inaccurate for the description of low-lying vibrational modes.[37] To overcome these limitations, different static approaches have been developed. They include the method introduced by Truhlar, referred here to as shifted RRHO (sRRHO), where all low frequency harmonic modes below a threshold are shifted to
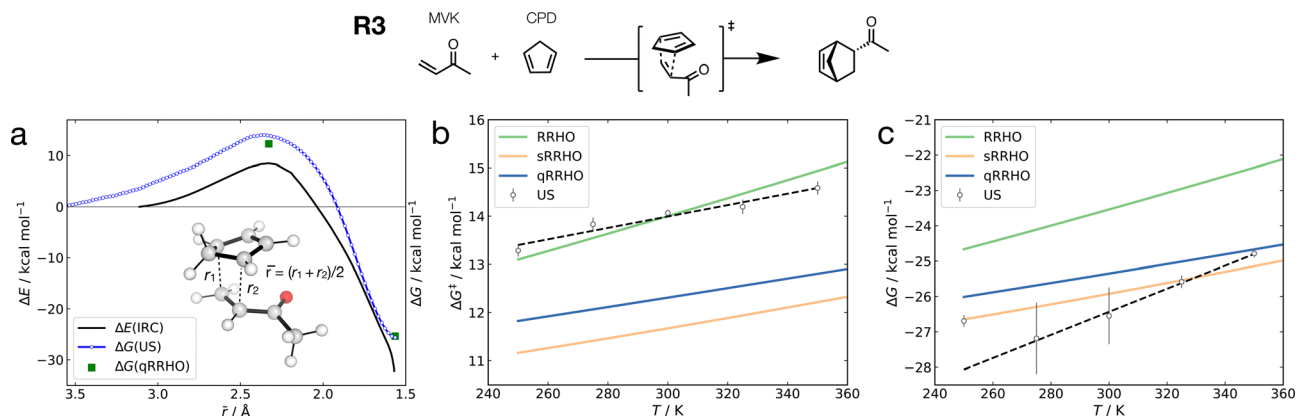
This journal is © the Owner Societies 2022

Phys. Chem. Chem. Phys., 2022, **24**, 20820–20827 | **20823**

Fig. 5 Reaction and activation free energies for the reaction between MVK and CPD, **R3**. (a) Comparison of the intrinsic reaction coordinate potential energy (ΔE), qRRHO free energy (ΔG_{qRROH}) and MLP-US free energy at 300 K ($\Delta G_{\text{MLP-US}}$). MLP-US was performed using the ACE MLP trained at the PBE0-D3BJ/def2-SVP level of theory at 500 K from the TS, then 30 windows were simulated over the reaction coordinate ($\bar{r} = (r_1 + r_2)/2$) for 10 ps. See Fig. S32 (ESI†) for umbrella histograms. (b) Activation free energy and (c) reaction free energy change as a function of temperature using different static endpoint methods and the MLP-US method. MLP-US values are averaged over 5 repeats and the error is the standard error of the mean.
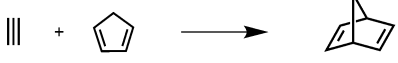
that value[38] and the qRRHO method introduced by Grimme, where those modes are treated between harmonic vibrations and rigid rotors.[37]

Using the ACE-MLPs trained for this reaction at the PBE0-D3BJ/def2-SVP level of theory we computed the free energy profile for **R3** using umbrella sampling (MLP-US) over an averaged reaction coordinate, $\bar{r} = (r_1 + r_2)/2$ (Fig. 5a). This approach enabled us to fully sample the different conformations not only at the critical points but along the reaction pathway, totalling 1.5 ns of sampling, which would be impossible to reach with AIMD. Moreover, our MLP-US simulations also capture anharmonic effects arising from low vibrations modes, which are not directly accounted for using the static methods. The MLP-US also provides a piece-wise free energy, where the choice of reaction coordinate is flexible.

Using this MLP, we also computed the temperature dependence for both activation and reaction energy in order to compare the MLP-US method to the static methods (Fig. 5b and c). For this DA reaction, MLP-US gives activation and reaction free energies comparable to the static methods, with a linear dependence on temperature. While no experimental data is available for the free energy of activation of this reaction in the gas phase, the comparison of the different static approaches to our MLP-US values suggest an estimate error of ≈ 3 kcal mol^{-1} for the free energy contributions. When comparing free energies of reaction (Fig. 5c), a larger deviation and larger error bars are observed at low temperature. This is likely due to the insufficient sampling; thus this point is excluded from the linear fit.

Using the same approach as for **R3**, we considered the DA reaction between cyclopentadiene and acetylene (**R4**, Table 1), for which comparison to experimental gas-phase kinetic data is possible.[39] Performing MLP-US enabled us to obtain quantitative accuracy in both enthalpic and entropic contributions. These results show that MLPs, in combination with US, enable us to achieve accurate free energies with associated errors, and

Table 1 Activation parameters for **R4** obtained from MLP-US. See ESI Section S10 for details. Standard error in the last digit is quoted in the parentheses

R4	$\Delta H^{\ddagger}$/kcal mol^{-1}	$\Delta S^{\ddagger}$/cal K^{-1} mol^{-1}
US	20(3)	−39(8)
Expt.[39]	21.9(1)	−37.3(2)

treat anharmonic effects without *ad hoc* corrections. While the MLP fitting and US protocol was completely automated and achieved within a day for this reaction, the choice of the QM reference method remains a critical choice that still requires human intervention. Comparison to a wider range of DA and other types of reactions, for which experimental data is available, will be the subject of future work.

3 Conclusion

In this work, we compared the performance of three MLP methods, ACE, NequIP and GAP, to study different aspects of modestly complex (≈ 50 atoms) Diels–Alder reactions. We also introduce a freely available automated fitting code *mlp-train* to facilitate the generation of these potentials.[40] While all three methods provide reasonable potentials, ACE and NequIP emerged as more efficient methods for generating accurate MLPs within 1 kcal mol^{-1} accuracy to the ground-truth surface, in particular for highly exothermic reactions.

ACE MLPs were also obtained for other Diels–Alder reactions and consistently achieved chemical accuracy (± 1 kcal mol^{-1}). This enabled us to perform RPMD in order to introduce nuclear quantum effects, as well as to obtain free energies employing umbrella sampling. While product distributions were strongly dependent on the DFT reference method, they were less

20824 | *Phys. Chem. Chem. Phys.*, 2022, **24**, 20820–20827

This journal is © the Owner Societies 2022

dependent on the type of dynamics propagated (classical *vs.* quantum). When computing free energies, MLP-US predicts barrier free energies within the static methods' ranges, which vary by ≈ 3 kcal mol^{-1}. For reaction free energies, larger errors bars were obtained at low temperature due to insufficient sampling. Comparison to experimental values demonstrates that MLP-US provide quantitative accuracy.

The diverse range of properties studied in this work demonstrate the applicability of our strategy, delivering accurate and efficient MLPs within a day. However, advancements in reference methods are needed to obtain accurate potentials without a preceding benchmark study. We are confident this will enable the routine development of chemically accurate reactive MLPs.

4 Methods

Training strategy

All Gaussian Approximation Potentials (GAPs) were trained using the *gap_fit* and *QUIP* codes with a Smooth Overlap of Atomic Positions (SOAP)[41] kernel with hyperparameters as defined in Table S2 (ESI†). The relationship between the size of the training set required to obtain an accurate GAP scales with system size is detailed for linear alkanes and small organic molecules in Section S5 (ESI†). GAP-MD simulations were performed with ASE[42] interfaced to *QUIP* with the *quippy* wrapper using the Langevin integrator. Initial configurations, CUR[30] selection and automated training set construction were generated with the *mlp-train* Python package (v. 1.0.0a).[40] The training set developed by AL was based on MLP-dynamics with the energy selection strategy 'diff', $E_T = 2.3$ kcal mol^{-1} (0.1 eV).[17] The other AL selection strategies, based on SOAP similarity, are discussed in Section S7 (ESI†).

A stable potential was defined by its ability to propagate 10×1 ps MD trajectories without finding a configuration $|E_{MLP} - E_0| > E_T$. Training sets for initial GAPs including those that used GP variance as a selection criteria were constructed using the *gap-train* code (v. 1.0.0b).[43] ACE[23,44] potentials were trained using the ACE.jl code[45] (v 0.8.4, Julia v. 1.6.3) and wrapped by pyjulip (v. 0.1, using PyCall). All MD simulations were propagated using the Langevin (0.02 friction coefficient) integrator in ASE v. 3.22.0 with initial velocities sampled from a Maxwell-Boltzmann distribution then propagated using a 0.5 fs timestep. NequIP[24] potentials used the nequip[46] v. 0.3.3 Python package (pytorch v. 1.9.1, pytorch_geometric v. 1.7.2, e2nn[47] v.0.3.5) with a maximum 1000 epochs and a 90:10 training to validation data split; all other hyperparameters were retained as their defaults (4.0 Å cut off radius, see nequip module in *mlp-train*, commit 0ba027c). All MLPs were trained on both energies and forces using the default loss functions for each MLP method with the default weights defined in the *mlp-train* package.[40] For GAP, the relative weighting on energies and forces is controlled by σ_E and σ_F, which values are reported in Table S2 (ESI†). ACE uses a square norm of the energy difference predicted and true forces (eqn (18) in ref. 44) with $w_E = 20$, $w_F = 1$. Finally, NequIP uses a square norm of the force

components (eqn (9) in ref. 24) with $w_E = 1$, $w_F = 100$. Gas phase (vacuum) simulations in periodic potentials used a cubic box ($l = 10$ nm). Combinations of harmonic restraining potentials were added and incorporated into the dynamics using ASE.

RPMD simulations

RPMD were performed using I-PI[48] with 16 beads, a 0.5 fs timestep, 300 K, initiated from the TS, or first equilibrated using constrained *NVT* dynamics as specified (80 fs, harmonic: 4 distances, $k = 1$ eV Å^{-2}) followed by *NVE* dynamics.

Ground truth calculations

DFTB calculations were performed with DFTB+[49] using 3ob[50] parameters in a periodic box with side lengths of 50 Å to mimic a vacuum. Molecular DFT used GFN2-XTB[51] in XTB v. 6.4.0. Molecular DFT (inc. AIMD), MP2 and coupled cluster [CCSD(T)] calculations performed with ORCA[52,53] v. 4.2.1 wrapped with autodE[54] using PBE,[55] PBE0,[56] B3LYP,[57] M06-2X[58] and ωB97X-D[59] functionals, def2-SVP, def2-TZVP, def2-TZVPP basis sets.[60] DLPNO coupled cluster calculations employed TightPNO thresholds.[61] Double-hybrid DFT calculations, B2PLYP,[62] DSD-PBEP86[63] and mPW2-PLYP[64] with def2-TZVP basis set, were performed in ORCA v. 5.0.2. Transition states were located with autodE. Plots were generated with matplotlib and, where quoted, include a standard error of the mean averaged over three independent repeats.

Author contributions

T. A. Y.: conceptualization, formal analysis, methodology, investigation, software, visualization, writing – original draft. T. J.-W.: conceptualization, formal analysis, investigation, software, visualization, writing – original draft. H. Z.: formal analysis, investigation, writing – review & editing. F. D.: conceptualization, analysis, funding acquisition, visualization, writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 A. J. Orr-Ewing, *Chem. Soc. Rev.*, 2017, **46**, 7597–7614.
2 J. Behler, *J. Chem. Phys.*, 2016, **145**, 170901.
3 M. Pinheiro, F. Ge, N. Ferré, P. O. Dral and M. Barbatti, *Chem. Sci.*, 2021, **12**, 14396–14413.

This journal is © the Owner Societies 2022

Phys. Chem. Chem. Phys., 2022, **24**, 20820–20827 | **20825**

4 A. P. Bartók, M. C. Payne, R. Kondor and G. Csányi, *Phys. Rev. Lett.*, 2010, **104**, 136403.

5 V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti and G. Csányi, *Chem. Rev.*, 2021, **121**, 10073–10141.

6 J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**, 146401.

7 H.-Y. Ko, J. Jia, B. Santra, X. Wu, R. Car and R. A. DiStasio Jr., *J. Chem. Theory Comput.*, 2020, **16**, 3757–3785.

8 S. Brickel, A. K. Das, O. T. Unke, H. T. Turan and M. Meuwly, *Electron. Struct.*, 2019, **1**, 024002.

9 J. Zeng, L. Cao, M. Xu, T. Zhu and J. Z. Zhang, *Nat. Commun.*, 2020, **11**, 1–9.

10 S. J. Ang, W. Wang, D. Schwalbe-Koda, S. Axelrod and R. Gómez-Bombarelli, *Chem*, 2021, **7**, 738–751.

11 M. Yang, L. Bonati, D. Polino and M. Parrinello, *Catal. Today*, 2022, **387**, 143–149.

12 M. de la Puente, R. David, A. Gomez and D. Laage, *J. Am. Chem. Soc.*, 2022, **144**, 10524–10529.

13 K. Töpfer, S. Käser and M. Meuwly, *Phys. Chem. Chem. Phys.*, 2022, **24**, 13869–13882.

14 J. Li and S. A. Lopez, *Acc. Chem. Res.*, 2022, **55**, 1972–1984.

15 J. Westermayr and P. Marquetand, *Chem. Rev.*, 2020, **121**, 9873–9926.

16 J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev and A. E. Roitberg, *J. Chem. Phys.*, 2018, **148**, 241733.

17 T. A. Young, T. Johnston-Wood, V. L. Deringer and F. Duarte, *Chem. Sci.*, 2021, **12**, 10944–10955.

18 A. M. Miksch, T. Morawietz, J. Kästner, A. Urban and N. Artrith, *Mach. Learn.*, 2021, **2**, 031001.

19 J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev and A. E. Roitberg, *Nat. Commun.*, 2019, **10**, 2903.

20 P. O. Dral, A. Owens, A. Dral and G. Csányi, *J. Chem. Phys.*, 2020, **152**, 204110.

21 J. Li and S. A. Lopez, *Chem. – Eur. J.*, 2022, **28**, e202200651.

22 Y. Zhao, N. González-García and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 2012–2018.

23 R. Drautz, *Phys. Rev. B*, 2019, **99**, 014104.

24 S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt and B. Kozinsky, *ArXiv*, 2021, **2101**, 03164.

25 K. Black, P. Liu, L. Xu, C. Doubleday and K. N. Houk, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 12860–12865.

26 W. J. Lording, T. Fallon, M. S. Sherburn and M. N. Paddon-Row, *Chem. Sci.*, 2020, **11**, 11915–11926.

27 M. Sato, S. Kishimoto, M. Yokoyama, C. S. Jamieson, K. Narita, N. Maeda, K. Hara, H. Hashimoto, Y. Tsunematsu, K. N. Houk, Y. Tang and K. Watanabe, *Nat. Catal.*, 2021, **4**, 223–232.

28 V. Martí-Centelles, A. L. Lawrence and P. J. Lusby, *J. Am. Chem. Soc.*, 2018, **140**, 2862–2868.

29 B. Briou, B. Améduri and B. Boutevin, *Chem. Soc. Rev.*, 2021, **50**, 11055–11097.

30 M. W. Mahoney and P. Drineas, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 697–702.

31 C. S. Jamieson, A. Sengupta and K. N. Houk, *J. Am. Chem. Soc.*, 2021, **143**, 3918–3926.

32 S. Habershon, D. E. Manolopoulos, T. E. Markland and T. F. Miller, *Annu. Rev. Phys. Chem.*, 2013, **64**, 387–413.

33 Q. Liu, L. Zhang, Y. Li and B. Jiang, *J. Phys. Chem. Lett.*, 2019, **10**, 7475–7481.

34 K. Black, P. Liu, L. Xu, C. Doubleday and K. N. Houk, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 12860–12865.

35 Z. Yang, C. Doubleday and K. N. Houk, *J. Chem. Theory Comput.*, 2015, **11**, 5606–5612.

36 R. D. Levine, *Molecular Reaction Dynamics*, Cambridge Univ Press, 2005, pp. 184–187.

37 S. Grimme, *Chem. – Eur. J.*, 2012, **18**, 9955–9964.

38 R. F. Ribeiro, A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2011, **115**, 14556–14562.

39 R. Walsh and J. M. Wells, *Int. J. Chem. Kinet.*, 1975, **7**, 319–329.

40 T. A. Young and T. Johnston-Wood, mlp-train, 2021, **https://github.com/duartegroup/mlp-train**.

41 A. P. Bartók, R. Kondor and G. Csányi, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**, 184115.

42 A. H. Larsen, J. J. 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. B. Jensen, J. Kermode, J. R. Kitchin, E. L. 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, *J. Phys.: Condens. Matter*, 2017, **29**, 273002.

43 T. A. Young and T. Johnston-Wood, gap-train, 2021, **https://github.com/t-young31/gap-train**.

44 D. P. Kovács, C. V. D. Oord, J. Kucera, A. E. Allen, D. J. Cole, C. Ortner and G. Csányi, *J. Chem. Theory Comput.*, 2021, **17**(12), 7696–7711.

45 C. Ortner, L. Zhang, A. Ross, M. Sachs and C. van der Oord, ACE.jl, **https://github.com/ACEsuit/ACE.jl**.

46 S. Batzner, A. Musaelian, L. Sun, A. Johansson, M. Geiger and T. Smidt, NequIP, 2021, **https://github.com/mir-group/nequip**.

47 M. Geiger, T. Smidt, M. Alby, B. K. Miller, W. Boomsma, B. Dice, K. Lapchevskyi, M. Weiler, M. Tyszkiewicz, S. Batzner, D. Madisetti, M. Uhrin, J. Frellsen, N. Jung, S. Sanborn, M. Wen, J. Rackers, M. Rød and M. Bailey, e3nn/e3nn: 2021-12-15, 2021, **https://zenodo.org/record/3724963**.

48 V. Kapil, M. Rossi, O. Marsalek, R. Petraglia, Y. Litman, T. Spura, B. Cheng, A. Cuzzocrea, R. H. Meißner, D. M. Wilkins, B. A. Helfrecht, P. Juda, S. P. Bienvenue, W. Fang, J. Kessler, I. Poltavsky, S. Vandenbrande, J. Wieme, C. Corminboeuf, T. D. Kühne, D. E. Manolopoulos, T. E. Markland, J. O. Richardson, A. Tkatchenko, G. A. Tribello, V. Van Speybroeck and M. Ceriotti, *Comput. Phys. Commun.*, 2019, **236**, 214–223.

49 B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M. Y. Deshaye, T. Dumitrică, A. Dominguez, S. Ehlert, M. Elstner, T. van der Heide, J. Hermann, S. Irle, J. J. Kranz, C. Köhler, T. Kowalczyk,

20826 | *Phys. Chem. Chem. Phys.*, 2022, **24**, 20820–20827

This journal is © the Owner Societies 2022

T. Kubař, I. S. Lee, V. Lutsker, R. J. Maurer, S. K. Min, I. Mitchell, C. Negre, T. A. Niehaus, A. M. N. Niklasson, A. J. Page, A. Pecchia, G. Penazzi, M. P. Persson, J. Řezáč, C. G. Sánchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V. W. Z. Yu and T. Frauenheim, *J. Chem. Phys.*, 2020, **152**, 124101.

50 M. Gaus, X. Lu, M. Elstner and Q. Cui, *J. Chem. Theory Comput.*, 2014, **10**, 1518–1537.

51 C. Bannwarth, S. Ehlert and S. Grimme, *J. Chem. Theory Comput.*, 2019, **15**, 1652–1671.

52 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2018, **8**, e1327.

53 F. Neese, F. Wennmohs, U. Becker and C. Riplinger, *J. Chem. Phys.*, 2020, **152**, 224108.

54 T. A. Young, J. J. Silcock, A. J. Sterling and F. Duarte, *Angew. Chem., Int. Ed.*, 2021, **60**, 4266–4274.

55 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

56 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.

57 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

58 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.

59 Y.-S. Lin, G.-D. Li, S.-P. Mao and J.-D. Chai, *J. Chem. Theory Comput.*, 2013, **9**, 263–272.

60 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297.

61 Y. Guo, C. Riplinger, U. Becker, D. G. Liakos, Y. Minenkov, L. Cavallo and F. Neese, *J. Chem. Phys.*, 2018, **148**, 011101.

62 S. Grimme, *J. Chem. Phys.*, 2006, **124**, 034108.

63 S. Kozuch and J. M. L. Martin, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20104–20107.

64 T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.*, 2006, **8**, 4398–4401.

This journal is © the Owner Societies 2022

Phys. Chem. Chem. Phys., 2022, **24**, 20820–20827 | **20827**