PERSPECTIVE • OPEN ACCESS

High-dimensional potential energy surfaces for molecular simulations: from empiricism to machine learning

View the article online for updates and enhancements.

MACHINE LEARNING
Science and Technology

PERSPECTIVE

High-dimensional potential energy surfaces for molecular simulations: from empiricism to machine learning

Oliver T Unke[1,3] , Debasish Koner[1] , Sarbani Patra[1], Silvan Käser[1] and Markus Meuwly[1,2]

[1] Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland
[2] Department of Chemistry, Brown University, Providence, RI, United States of America
[3] Present Address: Machine Learning Group, TU Berlin, Marchstr. 23, D-10587, Berlin, Germany.

E-mail: m.meuwly@unibas.ch

Keywords: potential energy surfaces, neural networks, reproducing kernel Hilbert space, chemical reactions

Abstract

An overview of computational methods to describe high-dimensional potential energy surfaces suitable for atomistic simulations is given. Particular emphasis is put on accuracy, computability, transferability and extensibility of the methods discussed. They include empirical force fields, representations based on reproducing kernels, using permutationally invariant polynomials, neural network-learned representations and combinations thereof. Future directions and potential improvements are discussed primarily from a practical, application-oriented perspective.

1. Introduction

The dynamics of molecular (i.e. chemical, biological and physical) processes is governed by the underlying intermolecular interactions. These processes can span a wide range of temporal and spatial scales and make a characterization and the understanding of elementary processes at an atomistic scale a formidable task [1]. Examples for such processes are chemical reactions or functional motions in proteins. For typical organic reactions the time scales are on the order of seconds whereas the actual chemical step (i.e. bond breaking or bond formation) occurs on the femtosecond time scale. In other words, during $\sim 10^{15}$ vibrational periods energy is redistributed in the system until sufficient energy has accumulated along the preferred 'progression coordinate' for the reaction to occur [2]. Similarly, the biological process of 'allostery' couples two (or multiple) spatially separated binding sites of a protein which is used to regulate the affinity of certain substrates to a protein, thereby controlling metabolism [3]. According to the conventional view of allostery, a conformational change of the protein (that might however be very small [4]) is the source of a signal, but other mechanisms have been proposed as well which are based exclusively on structural dynamics [5]. Here, binding of a ligand at a so-called *allosteric site* increases (or decreases) the affinity for a substrate at a distant *active site*, and the process can span multiple time and spatial scales to the extent of the size of the protein itself. Hence, an allosteric protein can be viewed as a 'transistor', and complicated feedback networks of many such switches ultimately make up a living cell [6]. As a third example, freezing and phase transitions in water are entirely governed by intermolecular interactions. Describing them at a sufficient level of detail has been found extremely challenging and a complete, quantitative understanding of the phase diagram or the structural dynamics of liquid water is still not available [7, 8].

High-dimensional energy functions also play important roles for applications in material sciences and catalysis. For example, the interaction between nanoparticles used as catalysts and their substrates can depend sensitively on the size and shape of the nanoparticle. Hence, an accurate description of the intermolecular interactions is required depending on the physical appearance of the particle [9] as well as for the examination of its dynamics [10]. In surface science, accurate potential energy surfaces (PESs) are, e.g. required to investigate the effect of substrate surface energy, orbital radii and ionization energy on monolayer metal oxide coating stability on support metal oxides [11]. Band gaps are another energetic property of materials that depends on the

composition and atomic ordering for which extensive information based on computation and experiment can provide deeper insight at a molecular level [12, 13]. Such efforts have potential applications in guiding the search for new photoactive materials for photocatalysis [14]. Finally, high dimensional PES are also probed when investigating the dynamics and energetics of phase transitions which has recently been done for melting points of a number of metalloids [15].

All the above situations require means to compute—for given nuclear coordinates of all particles involved—the total energy of the system efficiently and accurately. The most accurate and comprehensive approach is to solve the electronic Schrödinger equation for every nuclear configuration $\vec{x}$ of the system for which energies and forces are required. However, there are certain limitations which are due to the computational approach *per se*, e.g. the speed and efficiency of the method, or due to practical aspects of quantum chemistry such as accounting for the basis set superposition error, the convergence of the Hartree–Fock wavefunction to the desired electronic state for arbitrary geometries, or the choice of a suitable active space irrespective of molecular geometry for problems with multi-reference character, to name a few. Improvements and future avenues for making approaches based on quantum mechanics (QM) even more broadly applicable have been recently discussed [16]. For problems that require extensive conformational sampling or sufficient statistics purely QM-based dynamics approaches are still impractical.

A promising use of QM-based methods are mixed QM/molecular mechanics (QM/MM) treatments which are particularly popular for biophysical and biochemical applications [17]. Here, the system is decomposed into a 'reactive region' which is treated with a quantum chemical (or semiempirical) method and an environment described by an empirical force field. Such a decomposition considerably speeds up simulations such that even free energy simulations in multiple dimensions can be computed [18]. One of the current open questions in such QM/MM simulations is that of the size of the QM region required for converged results [19].

Other possibilities to provide energies for molecular systems are based on empirical energy expressions, fits of reference energies to reference data from quantum chemical calculations, representations of the energies by kernels or by using neural networks. These methods are the topic of the present perspective as they have shown to provide means to follow the dynamics of molecular systems over long time scales or to allow statistically significant sampling of the process of interest.

First, explicit representations of energy functions are discussed. This usually requires one to choose a functional form of the model function. Next, machine learned PESs are discussed. In a second part, topical applications and an outlook for these methods are presented.

2. Explicit representations

Empirical force fields (FFs) are one of the most seasoned concepts to represent the total energy of a molecular system given the coordinates $\vec{x}$ of all atoms. A general expression for an empirical FF includes bonded (E_{bonded}) and nonbonded ($E_{\text{nonbonded}}$) terms.

$$
E(\mathbf{x}) = \sum_{\text{bonds}} k_b (r - r_e)^2 + \sum_{\text{angle}} k_\theta (\theta - \theta_e)^2 \sum_{\text{dihedrals}} k_\chi (1 + \cos n\chi - \delta)
$$

$$
+ \sum_{\text{impropers}} k_{\text{imp}} (\phi - \phi_0)^2 + \sum_{\text{nonbonded}} \epsilon_{ij} \left[\left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{12} - \left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{12} \right] + \frac{q_i q_j}{\epsilon_l r_{ij}}. \tag{1}
$$

Such representations can be evaluated very efficiently, the forces are readily available and systems containing millions of atoms can be simulated for extended time scales [20]. On the other hand, the quantitative accuracy of such force fields as compared with high-level electronic structure methods is very limited. Conversely, one of the noteworthy advantages of empirical energy functions is that they can be consistently improved, for example by replacing harmonic potentials for chemical bonds by Morse oscillator functions or by extending conventional point charge electrostatics through multipolar series expansions [21–25]. Also, additional terms can be included to provide a more physically motivated representation, such as adding terms for polarization interactions [26].

While equation (1) is a general form of a FF for biomolecular applications, alternative functional forms and other applications have also been discussed in the literature. One example is the universal force field for which the parameters are based only on the element, its hybridization, and its connectivity [27]. It has been applied to organic, main group inorganic and transition metal-containing compounds. Another example is COMPASS (condensed-phase optimized molecular potentials for atomistic simulation studies) which has been applied to study organic molecules, inorganic small molecules, and polymers [28]. It relies heavily on fitting to reference data from electronic structure calculations but also includes refinement with respect to experimental data in the condensed phase. As a third example, DREIDING is a simple generic force field for predicting structures and dynamics of organic, biological, and main-group inorganic molecules [29]. Finally, there are also FFs that are particularly suitable to treat systems including transition metals or delocalized electronic structure. One of them

is based on valence bond concepts (VALBOND) [30, 31] which has also been extended to treat electronic effects such as the trans influence [32] or reactions [33].

For smaller molecular systems more accurate representations are possible. Typically, reference energies are computed from quantum chemical calculations on a grid (regular or irregular) of molecular geometries. These energies are then fit to parameters in a predetermined functional form to minimize the difference between the reference energies and the model function.

One example for such a predefined functional form are permutationally invariant polynomials (PIPs) which have been applied to molecules with 4–10 atoms and to investigate diverse physico-chemical problems [34]. Using PIPs, the permutational symmetry arising in many molecular systems is explicitly built into the construction of the parametrized form of the PES. The monomials are of the form $y_{ij} = \exp{(-r_{ij}/a)}$ where the r_{ij} are atom–atom separations and a is a range parameter. The total potential is then expanded into multinomials, i.e. products of monomials with suitable expansion coefficients. For an A_2B molecule the symmetrized basis which explicitly obeys permutational symmetry is $y_{12}^a(y_{13}^b y_{23}^c + y_{23}^b y_{13}^c)$. A library for constructing the necessary polynomial basis has been made publicly available [35].

One application of PIPs includes the dissociation reaction of CH_5^+ to $CH_3^+ + H_2$ for which more than 36 000 energies [36] were fitted with an accuracy of 78.1 cm^{-1}. With this PES the branching ratio to form HD and H_2 for CH_4D^+ and CH_5^+, respectively, was determined. Also, the infrared spectra of various isotopes were computed with this PES [37]. Other applications concern a fitted energy function for water dimer [38], which became the basis for the WHBB force field for liquid water [39] and that for acetaldehyde [40]. For acetaldehyde roughly 135 000 energies at the CCSD(T)/cc-pVTZ level of theory were fitted to 2655 terms with order 5 in the polynomial basis and 9953 terms with order 6 in the polynomial basis. For the relevant stationary states in that study the difference between the reference calculations and the fit ranges from 2.0 to 4.5 kcal mol^{-1}. However, the overall RMSD for all fitted points has not been reported [40]. With this PES the fragment population for dissociation into $CH_3 + HCO$ and $CH_4 + CO$ was investigated.

Another fruitful approach are double many body expansions [41]. These decompose the total energy of a molecular system first into one- and several many-body terms and then represent each of them as a sum of short- and long-range contributions [41]. This yields, for example, an RMSD of 0.99 kcal mol^{-1} for 3701 fitted points from electronic structure calculations at the multi reference configuration interaction (MRCI) level of theory for CNO [42]. As a comparison, another recent investigation of the same system [43] using a reproducing kernel Hilbert space (RKHS, see further below) representation yielded an RMSD of 0.38, 0.48 and 0.47 kcal mol^{-1} for the $^2A'$, $^2A''$ and $^4A''$ electronic states using more than 10 000 *ab initio* points for each surface.
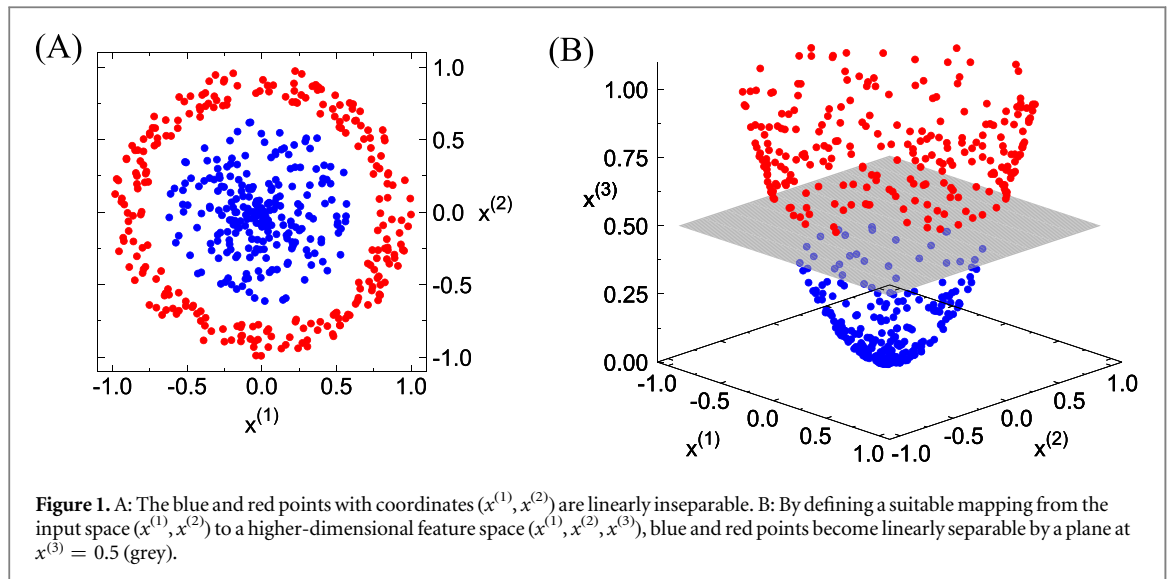
Local interpolation has also been shown to provide a meaningful approach. One such method is Shepard interpolation which represents the PES as a weighted sum of FFs, expanded around several reference geometries [44, 45]. Also, recently several computational resources have been made available to construct fully-dimensional PESs for polyatomic molecules such as Autosurf [46] or a repository to automatically construct PIPs.

Empirical FFs or those based on an RKHS representation can also be mixed to investigate chemical reactions. Because traditionally, empirical force fields are designed for one connectivity, they are not *a priori* suitable for studies of chemical reactions (bond breaking and bond formation). Several approaches have been devised in the past, including multi-state adiabatic reactive MD (MS-ARMD) [47], time-based reactive MD [48], or empirical valence bond theory (EVB) [49]. They all rely on mixing PESs for different states which provides the means to change from one bonding pattern to another one in a continuous fashion.

3. Machine learned PESs

Machine learning (ML) methods have become increasingly popular in recent years for constructing PESs, or estimate other properties of unknown compounds or structures [50–53]. Such approaches give computers the ability to learn patterns in data without being explicitly programmed [54], i.e. it is not necessary to complement a ML model with chemical knowledge. For example, no pre-conceived notion of bonding patterns needs to be assumed. For PES construction, suitable reference data are e.g. energy, forces, or both, usually obtained from *ab initio* calculations. Contrary to the explicit representations discussed in section 2, ML-based PESs are non-parametric and not limited to a predetermined functional form.

Most ML methods used for PES construction are either kernel-based or rely on artificial neural networks (NNs). Both variants take advantage of the fact that many nonlinear problems, such as predicting energy from nuclear positions, can be linearised by mapping the input to a (often higher-dimensional) feature space (see figure 1) [55]. Kernel-based methods utilize the *kernel trick* [56–58], which allows to operate in an implicit feature space without explicitly computing the coordinates of data in that space (see section 3.1 for more details). ML methods based on artificial NNs rely on 'neuron layers', which map their input to feature spaces by linear transformations with learnable parameters, followed by a nonlinearity called 'activation function'. Often, many

Figure 1. A: The blue and red points with coordinates $(x^{(1)}, x^{(2)})$ are linearly inseparable. B: By defining a suitable mapping from the input space $(x^{(1)}, x^{(2)})$ to a higher-dimensional feature space $(x^{(1)}, x^{(2)}, x^{(3)})$, blue and red points become linearly separable by a plane at $x^{(3)} = 0.5$ (grey).

such layers are stacked on top of each other to build increasingly complex feature spaces (see section 3.2). In the following, both variants are discussed in more detail.

3.1. Reproducing kernel representations

Starting from a data set $\{(y_i; \mathbf{x}_i)\}_{i=1}^N$ of N observations $y_i \in \mathbb{R}$ given the input $\mathbf{x}_i \in \mathbb{R}^D$, kernel regression aims to estimate unknown values y_* for input $\mathbf{x}_*$. For a PES, y is typically the total interaction energy and $\mathbf{x}$ is a representation of chemical structure (i.e. a vector of internal coordinates, a molecular descriptor like the Coulomb matrix [50], descriptors for atomic environments, e.g. symmetry functions [59], SOAP [60] or FCHL [61, 62], or a representation of crystal structure [63–65]). The representer theorem [66] for a functional relation $y = f(\mathbf{x})$ states that $f(\mathbf{x})$ can always be approximated as a linear combination

$$f(\mathbf{x}) \approx \widetilde{f}(\mathbf{x}) = \sum_{i=1}^N \alpha_i K(\mathbf{x}, \mathbf{x}_i), \tag{2}$$

where α_i are coefficients and $K(\mathbf{x}, \mathbf{x}')$ is a kernel function. A function $K(\mathbf{x}, \mathbf{x}')$ is a reproducing kernel of a Hilbert space $\mathcal{H}$ if the inner product $\langle \phi(\mathbf{x}), \phi(\mathbf{x}') \rangle$ of $\mathcal{H}$ can be expressed as $K(\mathbf{x}, \mathbf{x}')$ [67]. Here, ϕ is a mapping from the input space $\mathbb{R}^D$ to $\mathcal{H}$, i.e. $\phi: \mathbb{R}^D \mapsto \mathcal{H}$. Many different kernel functions are possible. Popular choices are the polynomial kernel

$$K(\mathbf{x}, \mathbf{x}') = \langle \mathbf{x}, \mathbf{x}' \rangle^d, \tag{3}$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product and d is the degree of the polynomial, or the Gaussian kernel given by

$$K(\mathbf{x}, \mathbf{x}') = e^{-\gamma \|\mathbf{x} - \mathbf{x}'\|^2} \tag{4}$$

with hyperparameter γ that determines the width of the Gaussian and $\|\cdot\|$ denotes the L^2-norm. It is also possible to include knowledge about the long range behaviour of the physical interactions into the kernel function itself [68] and the consequences of such choices on the long- and short-range behaviour of the inter- and extrapolation have been investigated in detail [69].

The mapping ϕ associated with the polynomial kernel (equation (3)) depends on the dimensionality of the input $\mathbf{x}$ and the chosen degree d of the kernel. For example, for $d = 2$ and two-dimensional input vectors, the mapping is $\phi: (x_1, x_2) \mapsto (x_1^2, \sqrt{2} x_1 x_2, x_2^2)$ and the Hilbert space $\mathcal{H}$ associated with the kernel function is three-dimensional. For a Gaussian kernel, $\mathcal{H}$ is even ∞-dimensional. This can easily be seen if equation (4) is rewritten as

$$K(\mathbf{x}, \mathbf{x}') = e^{-\gamma \|\mathbf{x}\|^2} e^{-\gamma \|\mathbf{x}'\|^2} e^{2\gamma \langle \mathbf{x}, \mathbf{x}' \rangle} \tag{5}$$

then the Taylor expansion of the third factor $e^{2\gamma \langle \mathbf{x}, \mathbf{x}' \rangle} = \sum_{d=0}^{\infty} \frac{1}{d!} (2\gamma \langle \mathbf{x}, \mathbf{x}' \rangle)^d$ reveals that the Gaussian kernel is equivalent to an infinite sum over polynomial kernels (scaled by constant terms). It is important to point out that in order to apply equation (2), the mapping ϕ has never to be calculated explicitly (or even known at all) and it is therefore possible to operate in the (high-dimensional) space $\mathcal{H}$ implicitly. This is often referred to as the *kernel trick* [56–58].

The coefficients α_i (equation (2)) can be determined such that $\widetilde{f}(\mathbf{x}_i) = y_i$ for all input $\mathbf{x}_i$ in the dataset, i.e.

$$\boldsymbol{\alpha} = \mathbf{K}^{-1}\mathbf{y}, \tag{6}$$

where $\boldsymbol{\alpha} = [\alpha_i \cdots \alpha_N]^{\mathrm{T}}$ is the vector of coefficients, $\mathbf{K}$ is an $N \times N$ matrix with entries $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ called kernel matrix [70, 71] and $\mathbf{y} = [y_1 \cdots y_N]^{\mathrm{T}}$ is a vector containing the N observations y_i in the data set. Since the kernel matrix is symmetric and positive-definite by construction, Cholesky decomposition [72] can be used to efficiently solve equation (6). Once the coefficients α_i have been determined, unknown values y_* at arbitrary positions $\mathbf{x}_*$ can be estimated as $y_* = \widetilde{f}(\mathbf{x}_*)$ using equation (2).

In practice however, the solution of equation (6) is only possible if the kernel matrix $\mathbf{K}$ is not ill-conditioned. Fortunately, in case $\mathbf{K}$ is ill-conditioned, a regularized solution can be obtained for example by Tikhonov regularization [73]. This amounts to adding a small positive constant λ to the diagonal of $\mathbf{K}$, such that

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda\mathbf{I})^{-1}\mathbf{y} \tag{7}$$

is solved instead of equation (6) when determining the coefficients α_i (here, $\mathbf{I}$ is the identity matrix). Adding $\lambda > 0$ to the diagonal of $\mathbf{K}$ damps the magnitude of the coefficients $\boldsymbol{\alpha}$ and increases the smoothness of $\widetilde{f}$. While this has the effect that the known values in the data set are only *approximately* reproduced by equation (2), i.e. strictly $\widetilde{f}(\mathbf{x}_i) \neq y_i$, perhaps counterintuitively, it can *increase* the overall quality of predictions for unknown $\mathbf{x}_*$: In cases where the values y_i are noisy, reproducing them exactly also reproduces the noise, which is unlikely to generalise to unknown data. Therefore, this method of determining the coefficients can also be used to prevent over-fitting and is known as kernel ridge regression (KRR).

KRR is closely related to Gaussian process regression (GPR) [74]. In GPR, it is assumed that the N observations $\{(y_i; \mathbf{x}_i)\}_{i=1}^{N}$ in the data set are generated by a Gaussian process, i.e. drawn from a multivariate Gaussian distribution with zero mean, and covariance $K(\mathbf{x}, \mathbf{x}')$. Note that a mean of zero can always be assumed without loss of generality since two multivariate Gaussian distributions with equal covariance matrix can always be transformed into each other by adding a constant term. Further, every observation y_i is considered to be related to $\mathbf{x}_i$ through an underlying function $f(\mathbf{x})$ and some observational noise (e.g. due to uncertainties in measuring y_i)

$$y_i = f(\mathbf{x}_i) + \mathcal{N}(0, \lambda), \tag{8}$$

where λ is the variance of the Gaussian noise model. The chosen covariance function $K(\mathbf{x}, \mathbf{x}')$ expresses an assumption about the nature of $f(\mathbf{x})$. For example, if the Gaussian kernel (equation (4)) is used, $f(\mathbf{x})$ is assumed to be smooth and the chosen Gaussian width γ determines how rapid $f(\mathbf{x})$ is allowed to change if the input $\mathbf{x}$ changes.

With these assumptions, it is now possible to determine the conditional probability $p(y_*|\mathbf{y})$, i.e. answer the question 'given the data $\mathbf{y} = [y_1 \cdots y_N]^{\mathrm{T}}$, how likely is it to observe the value y_* for an input $\mathbf{x}_*$?'. Since it was assumed that the data was drawn from a multivariate Gaussian distribution, it is possible to write

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathbf{K} + \lambda\mathbf{I} & \mathbf{K}_*^{\mathrm{T}} \\ \mathbf{K}_* & K(\mathbf{x}_*, \mathbf{x}_*) \end{bmatrix}\right), \tag{9}$$

where $\mathbf{K}$ is the kernel matrix (see equation (6)) and $\mathbf{K}_* = [K(\mathbf{x}_*, \mathbf{x}_1) \cdots K(\mathbf{x}_*, \mathbf{x}_N)]$. Then, the best (most likely) estimate for y_* is the mean of this distribution

$$\bar{y}_* = \mathbf{K}_*(\mathbf{K} + \lambda\mathbf{I})^{-1}\mathbf{y}. \tag{10}$$

Thus, estimating y_* with GPR (equation (10)) is mathematically equivalent to estimating y_* with KRR (compare to equations (2) and (7)). However, while in KRR, λ is only a hyperparameter related to regularization, in GPR, λ is directly related to the magnitude of the assumed observational noise (see equation (8)). Further, the predictive variance,

$$\mathrm{var}(y_*) = K(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{K}_*(\mathbf{K} + \lambda\mathbf{I})^{-1}\mathbf{K}_*^{\mathrm{T}} \tag{11}$$

which can also be derived from equation (9), can be useful to estimate the uncertainty of a prediction y_*, i.e. how confident the model is that its prediction is correct. Since KRR and GPR are so similar, they are both referred to as reproducing kernel representations in this work.

3.2. Artificial neural networks

The fundamental building blocks of artificial NNs [75–81] are so-called 'dense (neuron) layers', which transform input vectors $\mathbf{x} \in \mathbb{R}^{n_{\mathrm{in}}}$ linearly to output vectors $\mathbf{y} \in \mathbb{R}^{n_{\mathrm{out}}}$ through

$$\mathbf{y} = \mathbf{W}\mathbf{x} + \mathbf{b}, \tag{12}$$

where the weights $\mathbf{W} \in \mathbb{R}^{n_{\mathrm{out}} \times n_{\mathrm{in}}}$ and biases $\mathbf{b} \in \mathbb{R}^{n_{\mathrm{out}}}$ are parameters, and n_{in} and n_{out} denote the dimensionality of input and output, respectively. A single dense layer can therefore only represent linear relations. To model

nonlinear relationships between input and output, at least two dense layers need to be combined with a nonlinear function σ (called activation function), i.e.

$$\mathbf{h} = \sigma(\mathbf{W_1 x} + \mathbf{b_1}) \tag{13}$$

$$\mathbf{y} = \mathbf{W_2 h} + \mathbf{b_2}. \tag{14}$$

Such an arrangement (equations (13) and (14)) has been proven to be a general function approximator, meaning that any mapping between input $\mathbf{x}$ and output $\mathbf{y}$ can be approximated to arbitrary precision, provided that the dimensionality of the so-called 'hidden layer' $\mathbf{h}$ is large enough [82, 83]. As such, NNs are a natural choice for representing a PES, i.e. a mapping from chemical structure to energy (for PES construction, the output $\mathbf{y}$ usually is one-dimensional and represents the energy).

While *shallow* NNs with a single hidden layer (see above) are in principle sufficient to solve any learning task, in practice, *deep* NNs with multiple hidden layers are exponentially more parameter-efficient [84]. In a deep NN, l hidden layers are stacked on top of each other,

$$\mathbf{h_1} = \sigma(\mathbf{W_1 x} + \mathbf{b_1})$$
$$\mathbf{h_2} = \sigma(\mathbf{W_2 h_1} + \mathbf{b_2})$$
$$\vdots$$
$$\mathbf{h_l} = \sigma(\mathbf{W_l h_{l-1}} + \mathbf{b_l})$$
$$\mathbf{y} = \mathbf{W_{l+1} h_l} + \mathbf{b_{l+1}} \tag{15}$$

mapping the input $\mathbf{x}$ to increasingly complex feature spaces, until the features $\mathbf{h}_l$ in the final layer are linearly related to the output $\mathbf{y}$. The parameters of the NN, i.e. the entries in the matrices $\mathbf{W}_l$ and vectors $\mathbf{b}_l$, are initialized randomly and then optimized, for example via gradient descent, to minimize a loss function that measures the difference between the output of the NN and a given set of training data. For example, the mean squared error (MSE) is a popular loss function for regression tasks.
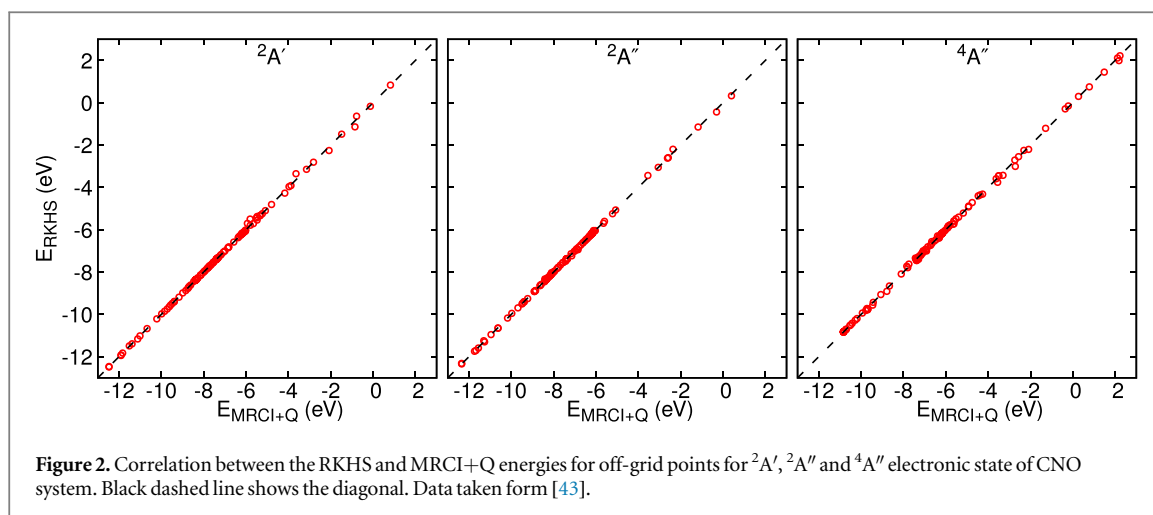
The earliest NN-based PESs directly use a set of internal coordinates, e.g. distances and angles, as input for the NN [85–89]. However, such approaches have the disadvantage that swapping symmetry equivalent atoms may also change the numerical values of the internal coordinates. Since it is not guaranteed that a NN maps two different inputs related by a permutation operation to the same output energy, the permutational invariance of the PES is violated. Another disadvantage of using internal coordinates as input is that a NN trained for a single molecule cannot be used to calculate the energy of a dimer, because they require a different number of internal coordinates for an unambiguous description of the molecular geometry. Therefore, for small systems, PESs based on NNs have been designed in the spirit of a many-body expansion [90–92], which circumvents these issues. However, these approaches involve a large number of individual NNs, i.e. one for each term in the many-body expansion and scale poorly for large systems.

For larger systems, it is common practice to decompose the total energy of a chemical system into atomic contributions, which are predicted by a single NN (or one for each element). This approach, known as high-dimensional neural network (HDNN) [93] and first proposed by Behler and Parrinello, relies on the chemically intuitive assumption that the contribution of an atom to the total energy depends mainly on its local environment.

Two variants of HDNNs can be distinguished: the 'descriptor-based' variant uses a hand-crafted descriptor [59, 94–96], to encode the environment of an atom, which is then used as input of a multi-layer feed-forward NN. Examples for this kind of approach are the 'Accurate NeurAl networK engINe for Molecular Energies' (ANAKIN-ME or ANI) [97] or TensorMol [98]. The 'message-passing' [99] variant directly uses nuclear charges and Cartesian coordinates as input and a deep neural network is used to exchange information ('messages') between individual atoms, such that a representation of their chemical environments is learned directly from the data. The deep tensor neural network [100] introduced by Schütt *et al* was the first NN of this kind and has since been refined in other NN architectures, for example SchNet [101], HIP-NN [102] or PhysNet [103]. Both types of HDNN perform well, however, the message-passing variant is able to automatically adapt the description of the chemical environments to the training data and the prediction task at hand and usually achieves a better performance [104].

4. Applications

In the following, illustrative applications of explicit representations of PESs (see section 2) and machine-learned PESs (see section 3) are discussed. PESs of sufficient quality for gas- and solution-phase reactions differ in at least two respects. While for reactions in the gas phase, typically involving small molecules as reactants, techniques to construct global, reactive PESs are becoming available, this is not so for reactions in solution. Often, the global property is also not required *a priori* for reactions in solution. Secondly, for reactions in the gas phase all

Figure 2. Correlation between the RKHS and MRCI+Q energies for off-grid points for $^2A'$, $^2A''$ and $^4A''$ electronic state of CNO system. Black dashed line shows the diagonal. Data taken form [43].

interactions are typically encoded in the global, reactive PES itself, whereas for reactions in solution the interaction between solute and solvent needs to be represented separately and explicitly. Therefore gas- and solution-phase are discussed in two different sections 4.1 and 4.2. While PESs are often used to explore the conformational space of a given system or study molecular (reaction) dynamics, machine-learned PESs can also serve as an alternative to *ab initio* methods for exploring chemical compound space. For example, it is possible to predict energies of molecules of different chemical composition from learning on a reference data set. Such applications are discussed briefly in section 4.3.
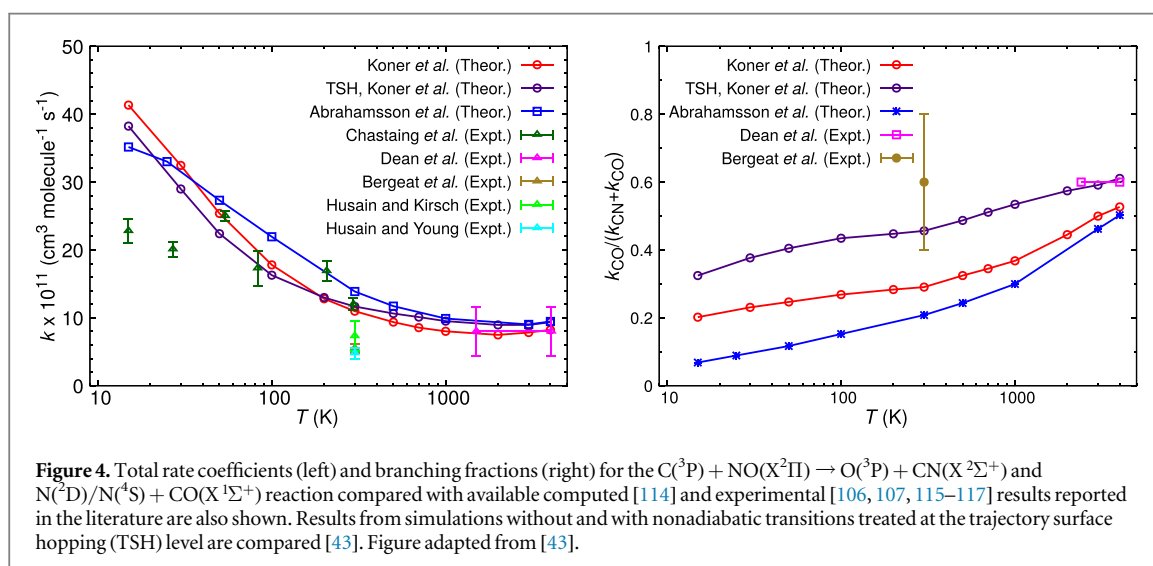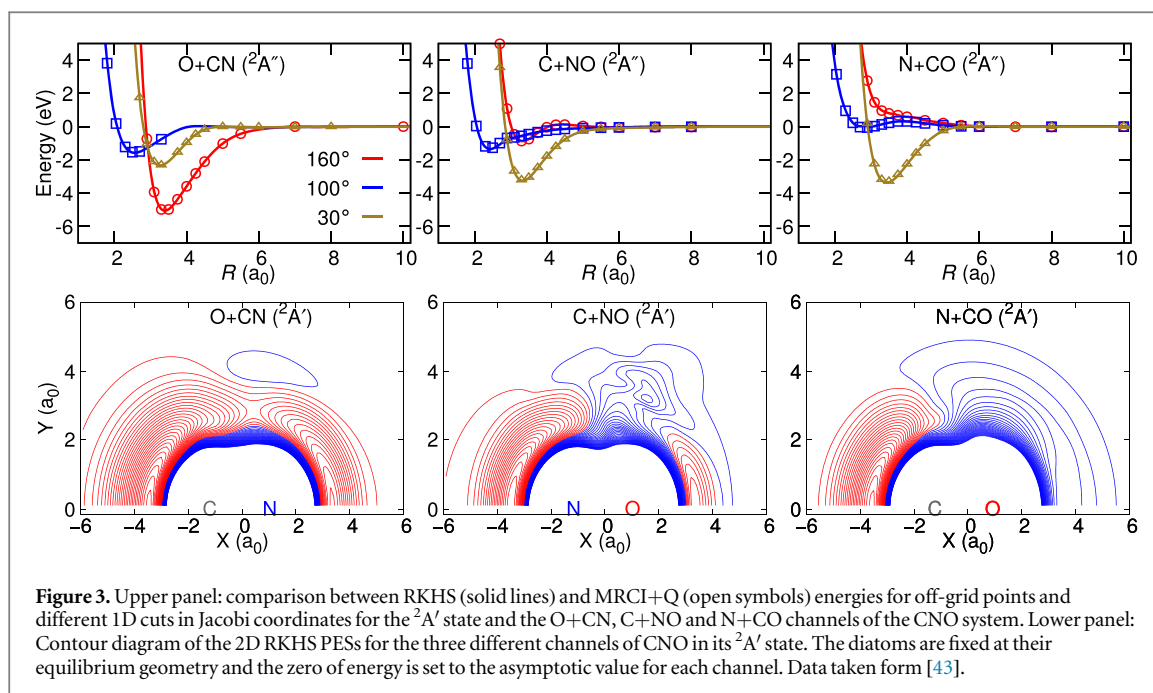
4.1. Gas phase reaction dynamics

4.1.1. Multisurface, reactive dynamics for triatomics

Triatomic systems constitute an important class of systems relevant to the chemistry in the atmosphere, in combustion and in the hypersonic regime upon reentry. Typical reactive collisions upon reentry of objects from outer space into Earth's atmosphere include the O+NO, O+CO, N+CO, C+NO, or N+NO reactions. Due to the high velocities of the impacting object, temperatures up to 20 000 K can be reached. To study the reaction dynamics at such high collision energies both, ground and lower electronically excited states need to be included. Hence, to describe the reactive dynamics for such systems, fully dimensional, reactive PESs including multiple electronic states are required. This is possible by using a large number of *ab initio* calculated energies at the MRCI level of theory and representing the PESs using a RKHS. Alternative approaches use explicit fitting of a parametrized form of a suitable many body expansion of the PES [41].

One example for such a system constitutes the reactive dynamics of [CNO] in the hypersonic regime at temperatures up to $T = 20\,000$ K [43]. The C+NO reaction is important in combustion chemistry and NO plays a crucial role in the chemistry near the surface of a space vehicle during atmospheric reentry [105]. For this, accurate fully dimensional PESs for the $^2A'$, $^2A''$ and $^4A''$ states were determined and used in quantum dynamics and quasiclassical trajectory (QCT) simulations. More than 50 000 *ab initio* energies were calculated at the MRCI+Q/aug-cc-pVTZ level of theory to construct the RKHS PESs. The electronic structure calculations were performed in grids based on Jacobi coordinates for each channel. RKHS was used to construct analytical representations for each channel and global 3D PESs were then generated by smoothly connecting the PESs for the three channels using switching functions. Correlation plots of the reference MRCI+Q and analytical energies obtained from the 3D RKHS based PESs are shown in figure 2 for the three electronic states to validate the quality of the RKHS-based global PESs. RKHS energies for different 1D cuts are compared with *ab initio* data in all three channels (C+NO, N+CO, and O+CN) for the $^2A''$ PES in figure 3 (top panel). The lower panel of figure 3 shows the topology of the $^2A'$ PES for all three channels. The overall good agreement between the *ab initio* and analytical energies in particular for the off-grid points in all channels and for all electronic states confirms the high quality of the PESs.

Experimental reference data is available for the rate coefficients and branching fractions of CO and CN products for the $C(^3P) + NO(X\,^2\Pi) \rightarrow O(^3P) + CN(X\,^2\Sigma^+)$ and $N(^2D)/N(^4S) + CO(X\,^1\Sigma^+)$ reaction [106–108]. From 40 000 QCT trajectories run at each temperature on each PES both, in an adiabatic and a nonadiabatic fashion within a Landau–Zener [109–112] formalism, the rate coefficients and branching fractions were determined. These rates can be directly compared with experiments and previous simulations. Figure 4 shows the rate coefficients and branching ratios for the products. Except for the lowest temperatures ($T \sim 30$ K and below) the rate coefficients agree well with experiments. This disagreement may be due to quantum effects or the fact that experiments for $T > 50$ K were carried out in Argon whereas for $T < 50$ K Helium was used as the buffer

Figure 3. Upper panel: comparison between RKHS (solid lines) and MRCI+Q (open symbols) energies for off-grid points and different 1D cuts in Jacobi coordinates for the $^2A'$ state and the O+CN, C+NO and N+CO channels of the CNO system. Lower panel: Contour diagram of the 2D RKHS PESs for the three different channels of CNO in its $^2A'$ state. The diatoms are fixed at their equilibrium geometry and the zero of energy is set to the asymptotic value for each channel. Data taken form [43].

Figure 4. Total rate coefficients (left) and branching fractions (right) for the $C(^3P) + NO(X^2\Pi) \rightarrow O(^3P) + CN(X^2\Sigma^+)$ and $N(^2D)/N(^4S) + CO(X^1\Sigma^+)$ reaction compared with available computed [114] and experimental [106, 107, 115–117] results reported in the literature are also shown. Results from simulations without and with nonadiabatic transitions treated at the trajectory surface hopping (TSH) level are compared [43]. Figure adapted from [43].

gas. Furthermore, it was found that including nonadiabatic transitions leads to better agreement with experiment within error bars but without nonadiabatic transitions the branching fractions were underestimated, see right panel in figure 4. In addition, computed final state distributions of the products for molecular beam-type simulations agree well with experiment. From such studies, thermal rates within an Arrhenius formalism can be determined which can then be used in more coarse grained simulations, such as direct simulation Monte Carlo (DSMC) [113].

4.1.2. Reactive dynamics of larger gas-phase systems

One recent application of multi state adiabatic reactive MD (MS-ARMD) and a NN-trained PES concerned the Diels-Alder reaction between 2,3-dibromo-1,3-butadiene (DBB) and maleic anhydride (MA) [118], see figure 5. DBB is a generic diene which fulfills the experimental requirements for conformational separation of its isomers by electrostatic deflection of a molecular beam [119, 120], thus enabling the characterization of conformational aspects and specificities of the reaction. MA is a widely used, activated dienophile which due to its symmetry simplifies the possible products of the reaction. The reaction of DBB and MA thus serves as a prototypical system well suited for the exploration of general mechanistic aspects of Diels–Alder reactions in the gas phase. The main questions concerned the synchronicity and concertedness of the reaction and how the reaction could be promoted. Until now, computational studies of Diels–Alder reactions including the molecular dynamics have
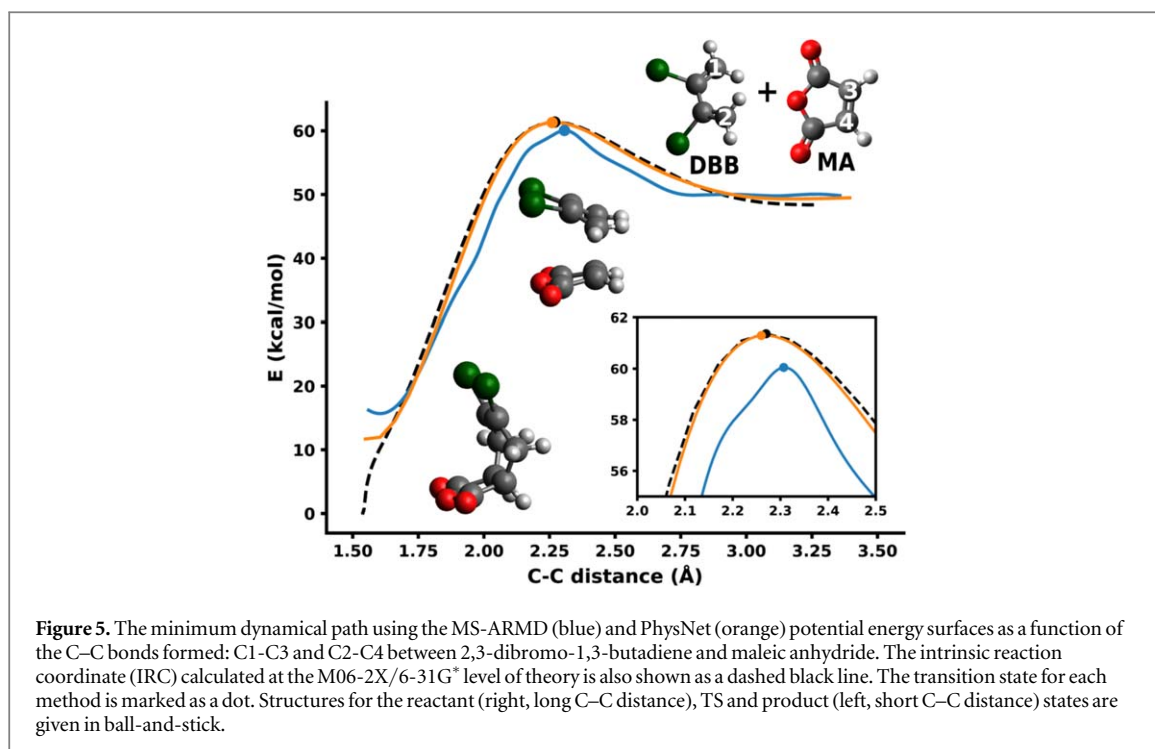
Figure 5. The minimum dynamical path using the MS-ARMD (blue) and PhysNet (orange) potential energy surfaces as a function of the C–C bonds formed: C1-C3 and C2-C4 between 2,3-dibromo-1,3-butadiene and maleic anhydride. The intrinsic reaction coordinate (IRC) calculated at the M06-2X/6-31G* level of theory is also shown as a dashed black line. The transition state for each method is marked as a dot. Structures for the reactant (right, long C–C distance), TS and product (left, short C–C distance) states are given in ball-and-stick.

started from transition state (TS)-like structures [121–125] or have used steered dynamics [126] both of which introduce biases and do not allow direct calculation of reaction rates.

To study the reaction in an unbiased fashion, two different reactive PESs were developed. One was based on the MS-ARMD approach whereas the second one employed the PhysNet [103] architecture to train a NN representation. For both representations scattering calculations were started from suitable initial conditions by sampling the internal degrees of freedom of the reactants and the collision parameter b. It is found that the majority of reactive collisions occur with rotational excitation and that most of them are synchronous. The relevance of rotational degrees of freedom to promote the reaction was also found when the minimum dynamical path [127] was calculated, see figure 5. The dynamics on both, the MS-ARMD and NN-trained PESs are very similar although the quality of the two surfaces is different. While the NN-trained PES is able to reproduce the training data to within 0.25 kcal mol^{-1} on average, the RMSD between reference and parametrized PES for MS-ARMD is 1.5 kcal mol^{-1} over a range of 80 kcal mol^{-1}. In terms of computational efficiency, MS-ARMD is, however, ∼200 times faster than PhysNet.

Another prototypical reaction scheme concerns S_N2 reactions. In a recent comparative study [128], three reactive PESs for the $[Cl–CH_3–Br]^-$ system were constructed: Two of these PESs rely on FFs, either combined with the MS-ARMD [47] or the MS-VALBOND [33] approach to construct the global reactive PES, whereas the third is NN-based. While all methods are able to fit the *ab initio* reference data with $R^2 > 0.99$, the NN-based PES achieves mean absolute and root mean squared deviations that are an order of magnitude lower than the other two methods when using the same number of reference data. When increasing the size of the reference data set, the prediction errors made by the NN-based PES are even up to three orders of magnitude lower than for the force field-based PESs. However, at the same time, evaluating the NN-based PES is about three orders of magnitude slower [128]. This comparative study demonstrates that different computational approaches are similarly suitable to investigate chemical reactions in the gas phase at an atomistic level. When considering the same reaction in solution, methods based on empirical force fields are probably still superior to more modern, ML-based PESs. This point is considered next.

4.2. Reactions in the condensed phase

For reactions in the condensed phase, two different situations are considered in the following. In one of them, ligands bind to a substrate anchored within a protein, such as for small diatomic ligands binding to the heme-group in globins. In the other, the substrate is chemically transformed as is the case for the Claisen rearrangement from chorismate to prephenate.

4.2.1. Ligand (Re-)binding in globins

Computationally, the structural dynamics accompanying NO-rebinding to Myoglobin has recently been investigated with the aim to assign the transient, metastable structures relevant for rebinding of the ligand on
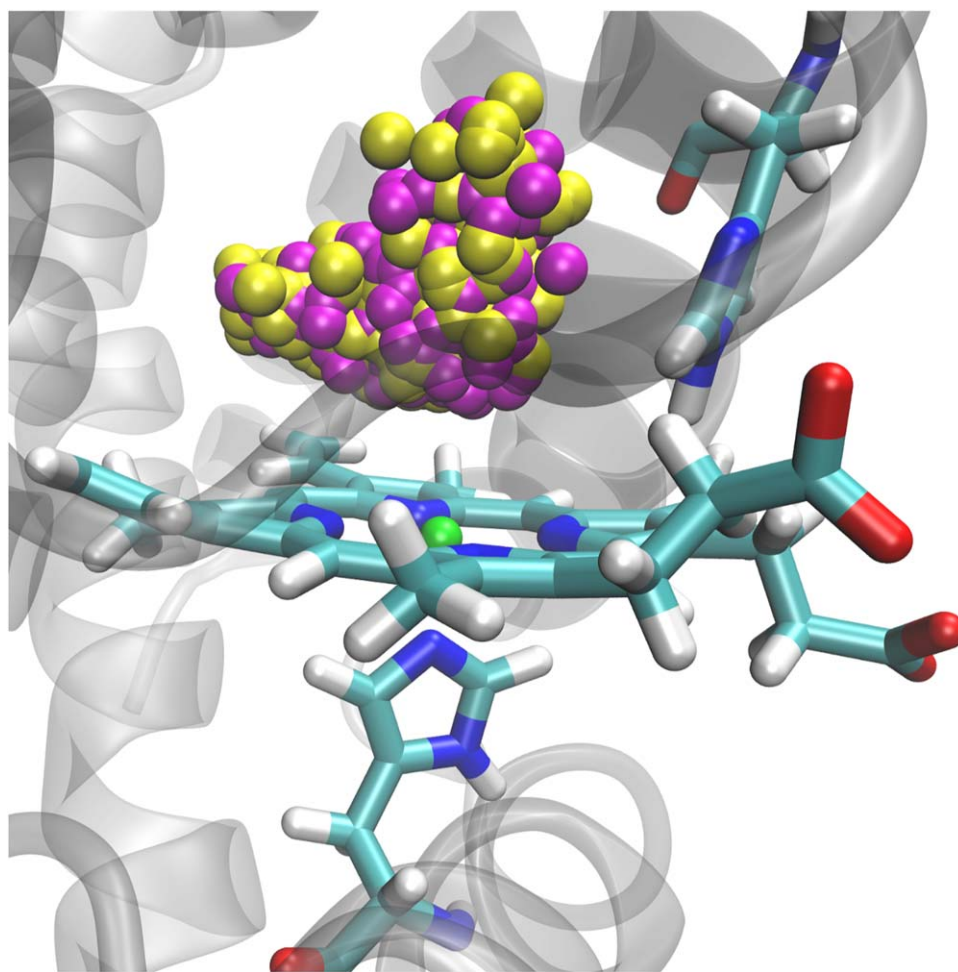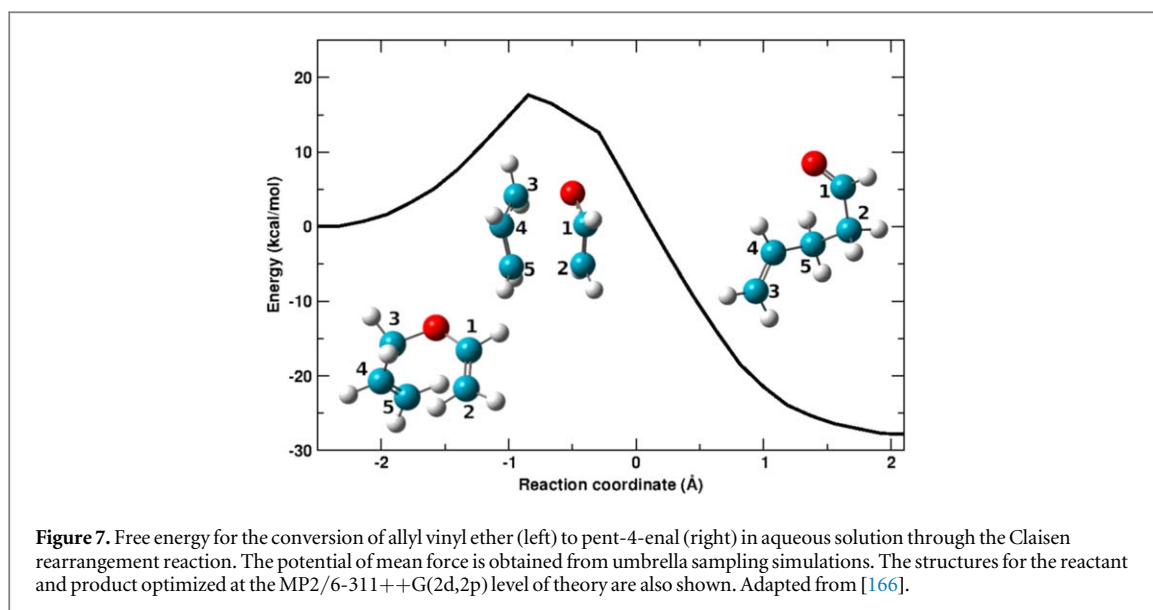
Figure 6. The active site of Myoglobin with the heme group (sceletal) and the rebinding NO ligand with the nitrogen in yellow and the oxygen in magenta, and the iron atom in green. The two histidine residues, His64 (to the right) and His93 (binding to the heme-iron from below) are also shown in sceletal representation. For His64 the 'out' conformation is shown. Adapted from [129].

different time scales [129]. For this, reactive MD simulations using MS-ARMD simulations were run involving the bound ^2A and the unbound ^4A states which are also probed experimentally. The energy for each of the states was represented as a reproducing kernel [68, 129, 130] for the subspace of important system coordinates (the heme(Fe)–NO separation and angle, and the doming coordinate of the heme-Fe) combined with an empirical force field for all remaining degrees of freedom, see figure 6 for the active structure of Mb. Such an approach is inspired by a decomposition of the system into a region that is modelled with high accuracy (typically a 'quantum region') and an environment (the 'molecular mechanics' part).

With a system parametrized in this fashion, extensive reactive MD simulations could be run [129]. The kinetics for ligand rebinding is nonexponential with time scales of 10 and 100 ps. These are consistent with time scales measured from optical, infrared, and x-ray absorption experiments and previous computational work [48, 131–141]. The influence of the iron-out-of-plane (Fe-oop or 'doming') coordinate on the rebinding reaction, as predicted by experiment [134], was directly established. The two time scales (10 and 100 ps) are associated with two structurally different states of the His64 side chain—one 'out' (state A, shown in figure 6) and one 'in' (state B)—which control ligand access and rebinding dynamics. Such an unequivocal assignment was not possible from experiment [142]. In addition, the simulations provide an explanation why an energetically feasible state for NO-binding to heme is typically not found in Mb: although the bound Fe-ON state is a local minimum on the PES, the energy of this state on the unbound ^4A manifold is lower and, hence, the bound ^2A Fe-ON can not be spectroscopically characterized. The simulations finally clarify that the x-ray absorption spectroscopy experiments are unable to distinguish between structures with photodissociated NO 'close to' or 'far away' from the heme-Fe in the active site as had been proposed [137].

In this fashion, validation of experimental results by the MD simulations and in-depth analysis of the configurations driving the dynamics on the different time scales (10 and 100 ps) allowed to identify the structural origins of the conformational dynamics at a molecular level, see figure 6. It is expected that further combined

Figure 7. Free energy for the conversion of allyl vinyl ether (left) to pent-4-enal (right) in aqueous solution through the Claisen rearrangement reaction. The potential of mean force is obtained from umbrella sampling simulations. The structures for the reactant and product optimized at the MP2/6-311++G(2d,2p) level of theory are also shown. Adapted from [166].

experimental and computational studies of this kind will provide the necessary insight to link energetics, structures and dynamics in complex systems.

4.2.2. Reactions in solution

The Claisen rearrangement [143] is an important [3,3]-sigmatropic rearrangement for high stereoselective [144] C–C bond formation [145]. The text book example of a Claisen rearrangement is the reaction of allyl-vinyl ether (AVE) to pent-4-enal [146], see figure 7. In polar solvent the stabilization of the TS relative to the reaction in vacuum is the origin of the catalytic effect [147–149]. This has motivated numerous studies on enzymatic Claisen rearrangements in particular [150–160] and reactions with related substrates [161–164]. Compared to the reaction in aqueous solution the enzymatic catalysis of the Claisen rearrangement reaction in chorismate mutase (CM) leads to a rate acceleration by $\sim 10^6$ due to stabilisation of the TS [165].

A reactive force field based on MS-ARMD was parametrized for AVE and used unchanged for AVE-$(CO_2)_2$ and chorismate to study their Claisen rearrangements in the gas phase, in water and in the chorismate mutase from *Bacillus subtilis* (BsCM) [166]. Using free energy simulations it is found that in going from AVE and AVE-$(CO_2)_2$ to chorismate and using the same reactive PES the rate slows down when going from water to the protein as the environment. A typical free energy profile for the conversion of AVE to pent-4-enal in water together with structures for the reactant, product and TS is shown in figure 7. For the largest substrate (chorismate) the simulations correctly find that the protein accelerates the reaction. Considering the changes of +4.6 (AVE), +2.9 (AVE-$(CO_2)_2$) and −4.4 (chorismate) kcal mol^{-1} in the activation free energies and correlating them with the actual chemical modifications suggests that both, electrostatic stabilization (AVE $\rightarrow$ AVE-$(CO_2)_2$) and entropic contributions (AVE-$(CO_2)_2 \rightarrow$ chorismate, through the rigidification and larger size of chorismate) lead to the rate enhancement observed for chorismate in CM.

As for the reaction itself it is found that for all substrates considered the O–C bond breaks prior to C–C bond formation. This agrees with kinetic isotope experiments according to which C–O cleavage always precedes C–C bond formation [167]. For the nonenzymatic thermal rearrangement of chorismate to prephenate the measured kinetic isotope effects [167, 168] indicate that at the TS the C–O bond is about 40% broken but little or no C–C bond is formed, consistent with an analysis based on 'More O'Ferrall-Jencks' (MOFJ) diagrams [169, 170].

The analysis of the TS position in the active site of BsCM reveals that the lack of catalytic effect on AVE is due to its loose positioning, insufficient interaction with and TS stabilization by the active site of the enzyme. Major contributions to localizing the substrate in the active site of BsCM originate from the CO_2^- groups. This together with the probability distributions in the reactant, TS and product states suggest that entropic factors must also be considered when interpreting differences between the systems, specifically (but not only) in the protein environment.

4.3. Energy predictions

The systematic exploration of chemical space is a possible way to find as of yet unknown compounds with specific properties, e.g. for medical or materials applications. For example, the GDB-17 database [171] enumerates 166 billion small organic molecules that are potential drug candidates. However, running *ab initio* calculations to determine the properties of billions of molecules is computationally infeasible. Machine-learned

PESs were shown to reach accuracies on par with hybrid DFT methods [172] and thus can serve as an efficient alternative to predict e.g. stabilization energy or equilibrium structures.

In order to be able to compare different approaches, benchmark datasets are used to assess the accuracy of ML methods. One of the most popular benchmarks for this purpose is QM9 [173]. It consists of several properties for 133 885 molecules in their minimum energy (equilibrium) geometries corresponding to a subset of all species with up to nine heavy atoms (C, O, N, and F) out of the GDB-17 database calculated at the B3LYP/ 6-31G(2df,p) level of theory [171]. For example, after training on 50 000 structures, both the PhysNet neural network architecture [103] and KRR based on the FCHL2019 descriptor [62] achieve a mean absolute error of ≈ 0.3 kcal mol^{-1} for predicting the energy of unseen molecules. When the FCHL2018 [61] descriptor is used in the kernel model, the same accuracy is reached after training on just 20 000 structures. However, FCHL2018 descriptors are computationally expensive and therefore difficult to apply to larger training set sizes [62].

It is also possible to predict other molecular properties (apart from energy) with ML methods. Interested readers are referred to the literature [172], where the accuracy of different approaches for the prediction of other properties, for example HOMO/LUMO energies, dipole moments, polarizabilities, zero point vibrational energies, or heat capacities has been compared. Since all molecular properties can be derived from the wave function, recent approaches aim to directly predict the electronic wave function from nuclear coordinates [174] or incorporate response operators into the model [175].

5. Outlook and conclusions

This section puts the methods discussed in the present overview into perspective and discusses future extensions together with their advantages and disadvantages of the approaches discussed so far.

Reproducing kernels have been applied to generate accurate representations of PES for different triatomic systems (3D) to study either reactive collisions or vibrational spectroscopy. The RKHS procedure can also be applied to construct higher dimensional PESs. As an example, an RKHS representation of the 6D PES for N$_4$ is discussed. Previously, a global PES was constructed for N$_4$ using PIPs from 16 435 CASPT2/maug-cc-pVTZ energies [176, 177] which are also used here. For constructing the RKHS, a total of 16 046 *ab initio* energies up to 1200 kcal mol^{-1} were used. The full PES is expanded in a many body expansion,

$$V(r_i) = \sum_{i=1}^{4} V_i^{(1)} + \sum_{i=1}^{6} V_i^{(2)}(r_i) + \sum_{i=1}^{3} V_i^{(3)}(r_j, r_k, r_l) + V_i^{(4)}(r_j, r_k, r_l, r_m), \tag{16}$$

where the first term is the sum of four 1-body energies, the second term is the sum of six 2-body interaction energies, the third term is the sum of four 3-body interaction energies and the last term is the 4-body interaction energy. The first term is a constant which is the dissociation energy of N$_4$ to four N atoms. Each 2-body term can be expressed by a 1D reproducing kernel polynomial and corresponding RKHS PESs can be constructed from the diatomic N$_2$ potential. The last two terms can be expressed by a product of three and six 1D reproducing kernel polynomials. In this work, the sum of the last two terms are calculated using RKHS interpolation of the $E^{(3+4)}$ energies. The sum of 3 and 4-body interaction energies ($E^{(3+4)}$) is calculated as

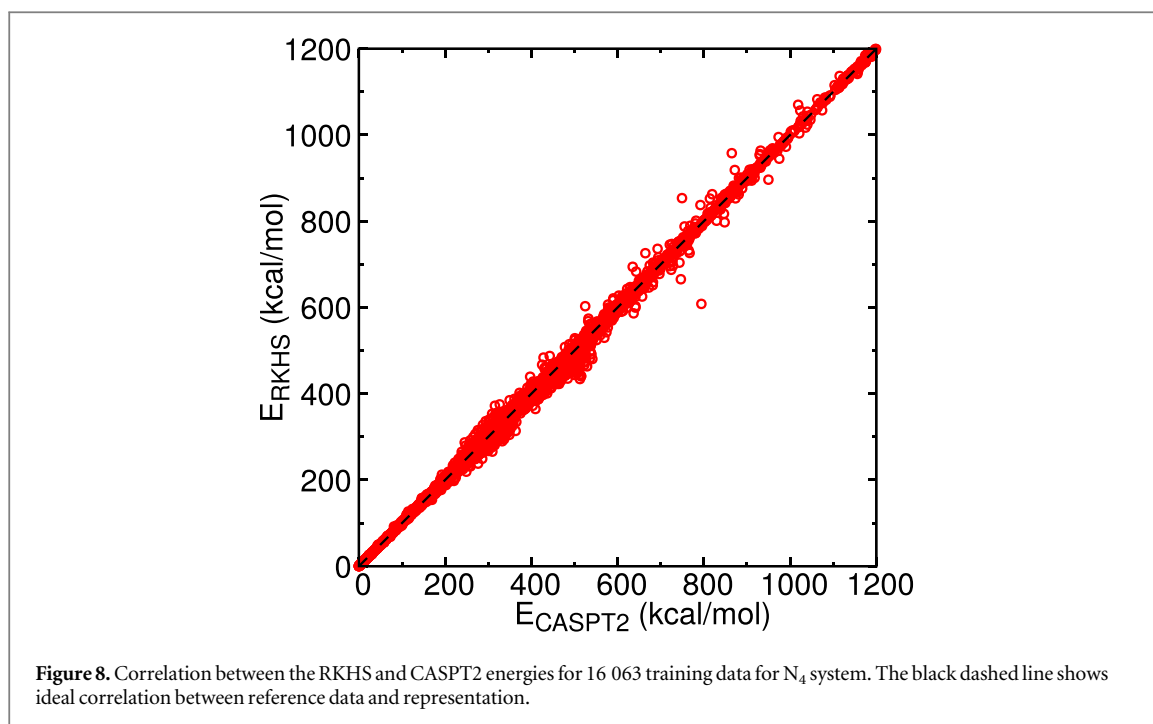$$E^{(3+4)} = V(r_i) - \sum_{i=1}^{4} V_i^{(1)} - \sum_{i=1}^{6} V_i^{(2)}(r_i). \tag{17}$$

For all the cases the 1D kernel function ($k^{n,m}$) with smoothness $n = 2$ and asymptotic decay $m = 6$ is used for the radial dimensions, which is expressed as

$$k^{2,6}(x, x') = \frac{1}{14} \frac{1}{x_>^7} - \frac{1}{18} \frac{x_<}{x_>^8}, \tag{18}$$

where, $x_>$ and $x_<$ are the larger and smaller values of x and x', respectively, and the kernel smoothly decays to zero at long range. Symmetry of the system can also be implemented within this approach by considering all possible combinations for the 3 and 4-body interaction energies. Here, it is worth to be mentioned that interpolating the 3-body and 4-body terms separately should provide more accurate energies, which is however not possible in this case as the triatomic energies are not available.

The root mean squared (RMSE) and mean absolute (MAE) errors are computed for the training data set and tabulated in table 1. The correlation between the reference *ab initio* energies and RKHS interpolated energies are reported in figure 8 with an $R^2 = 0.9981$. Elected potential energy curves for fixed bondlength r$_A$ of one N$_2$ molecule depending on the bondlength r$_B$ of the second N$_2$ and for different N$_2$–N$_2$ separations d and relative orientation are shown in 9. The *ab initio* energies in figure 9 are not included in constructing the RKHS representation and demonstrate that a RKHS can successfully reproduce the overall shape and values of the unknown *ab initio* potential.

Figure 8. Correlation between the RKHS and CASPT2 energies for 16 063 training data for N_4 system. The black dashed line shows ideal correlation between reference data and representation.

Table 1. Root mean squared and mean absolute errors computed for the training data from the RKHS-based PES in different energy ranges for N_4. Units of energies are in kcal mol^{-1}.
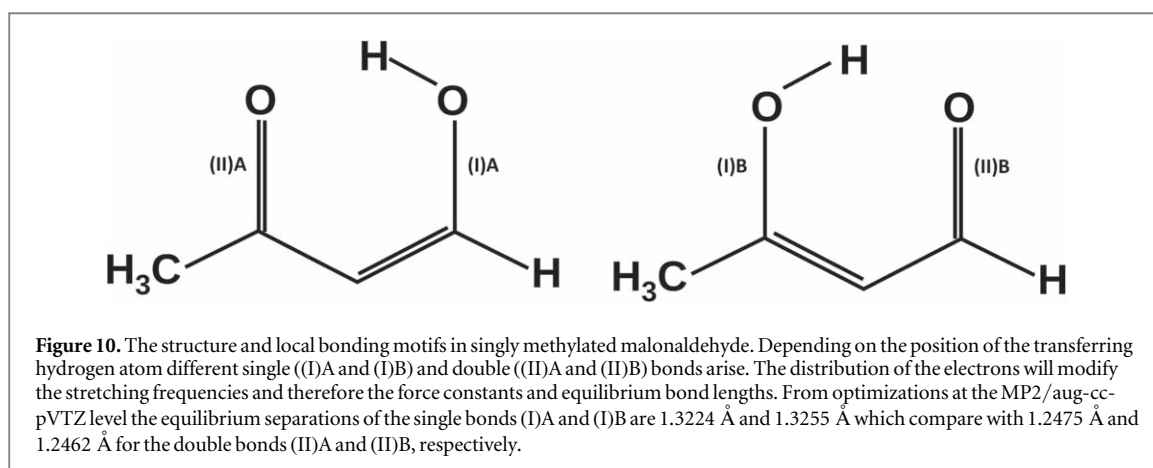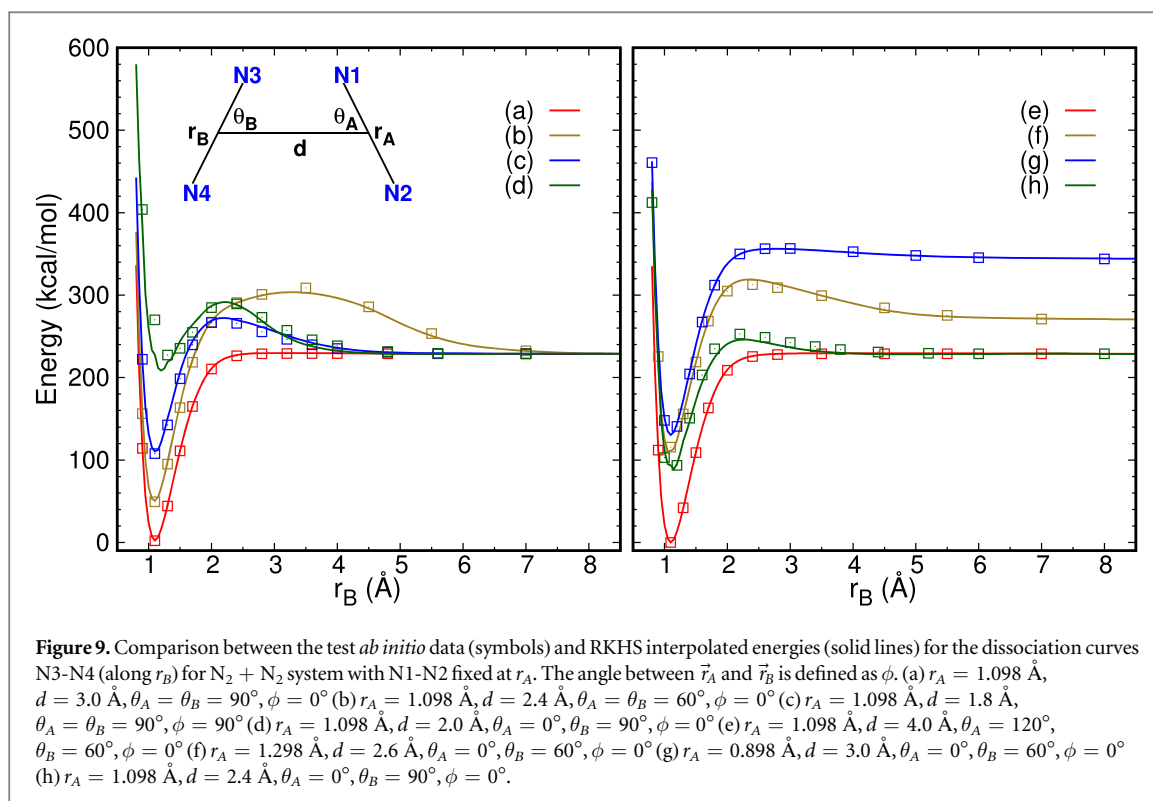
Energy range	Number of points	RMSE	MAE	RMSE [176]
$E \leqslant 100.0$	678	1.4	0.8	1.8
$100.0 < E \leqslant 228.0$	1894	3.3	1.8	4.1
$228.0 < E \leqslant 456.0$	11 707	6.9	3.6	7.2
$456.0 < E \leqslant 1000.0$	1608	16.5	9.6	18.0
$1000.0 < E < 1200.0$	159	9.1	4.8	

Although techniques such as RKHS or PIPs can provide accurate representations, their extensions to higher dimensions remains a challenge. Recently, the use of PIPs was demonstrated for the PES of N-methyl acetamide which is an important step in this direction [178]. Additionally, the (symmetric) Gradient Domain Machine Learning ((s)GDML) approach [179, 180] has been used to construct PESs for several small organic molecules, such as ethanol, malondialdehyde and aspirin [181]. Another challenge is to reduce the number of points required to define such a PES. Efforts in this direction have recently shown that with as few as 300 reference points the PES for scattering calculations in OH+H_2 can be described from a fit based on Gaussian processes together with Bayesian optimization [182]. Nevertheless, such high-accuracy representations of PESs for extended systems will remain a challenge for both, the number of high-quality reference calculations required and the type of inter- (and extra-)polation used to represent them.

Another important aspect of accurate studies of the energetics and dynamics of molecular systems concerns the observation, that 'chemistry' is often local. For example, the details of a chemical bond—its equilibrium separation and its strength—can depend sensitively on the local environment which may play an important role in applications such as infrared spectroscopy, as illustrated for methylated malonaldehyde, see figure 10. Depending on the position of the proton the nature of the CO bond changes. Overall, there are four chemically different CO bonds, two single bonds (I)A and (I)B, and two double bonds (II)A and (II)B. In the language of an empirical force field, the equilibrium bond lengths and the force constants change between these two structures in a dynamical fashion, depending on the position of the transferring hydrogen atom. Capturing such effects within an empirical FF is possible, but laborious, as was recently done for the oxalate anion [183].

Capturing such effects within a NN-trained global PES using PhysNet is more convenient. As an example, the situation in singly-methylated malonaldehyde (acetoacetaldehyde, AAA) is considered, see figure 10. There are two CO motifs each of which can carry the transferring hydrogen atom at the oxygen atom. Depending on whether the hydrogen atom is at the OC–CH_3 or OC–H side the chemical nature of the CO bond changes. This

Figure 9. Comparison between the test *ab initio* data (symbols) and RKHS interpolated energies (solid lines) for the dissociation curves N3-N4 (along r_B) for $N_2 + N_2$ system with N1-N2 fixed at r_A. The angle between $\vec{r}_A$ and $\vec{r}_B$ is defined as ϕ. (a) $r_A = 1.098$ Å, $d = 3.0$ Å, $\theta_A = \theta_B = 90°$, $\phi = 0°$ (b) $r_A = 1.098$ Å, $d = 2.4$ Å, $\theta_A = \theta_B = 60°$, $\phi = 0°$ (c) $r_A = 1.098$ Å, $d = 1.8$ Å, $\theta_A = \theta_B = 90°$, $\phi = 90°$ (d) $r_A = 1.098$ Å, $d = 2.0$ Å, $\theta_A = 0°$, $\theta_B = 90°$, $\phi = 0°$ (e) $r_A = 1.098$ Å, $d = 4.0$ Å, $\theta_A = 120°$, $\theta_B = 60°$, $\phi = 0°$ (f) $r_A = 1.298$ Å, $d = 2.6$ Å, $\theta_A = 0°$, $\theta_B = 60°$, $\phi = 0°$ (g) $r_A = 0.898$ Å, $d = 3.0$ Å, $\theta_A = 0°$, $\theta_B = 60°$, $\phi = 0°$ (h) $r_A = 1.098$ Å, $d = 2.4$ Å, $\theta_A = 0°$, $\theta_B = 90°$, $\phi = 0°$.

Figure 10. The structure and local bonding motifs in singly methylated malonaldehyde. Depending on the position of the transferring hydrogen atom different single ((I)A and (I)B) and double ((II)A and (II)B) bonds arise. The distribution of the electrons will modify the stretching frequencies and therefore the force constants and equilibrium bond lengths. From optimizations at the MP2/aug-cc-pVTZ level the equilibrium separations of the single bonds (I)A and (I)B are 1.3224 Å and 1.3255 Å which compare with 1.2475 Å and 1.2462 Å for the double bonds (II)A and (II)B, respectively.

also influences the frequencies of the CO stretch vibrations. Figure 11 reports the infrared spectrum from normal modes from MP2/aug-cc-pVTZ calculations and from an NN trained on energies, forces and dipole moments at the same level of theory. As is shown, the normal modes from the electronic structure calculations from the MP2/aug-cc-pVTZ for the two isomers (top and bottom panels) differ appreciably in the range of the amide-I stretches. Above 1600 cm^{-1} the harmonic frequencies occur at 1644 and 1692 cm^{-1} for isomer AAA1 and at 1658 and 1696 cm^{-1} for isomer AAA2. The NN (middle two panels) is successful in capturing the higher frequency (at 1689 and 1695 cm^{-1} for the two isomers, respectively) whereas for the lower frequency the two modes occur at 1635 and 1634 cm^{-1}. Additional modes involving CO stretch vibrations occur between 1400 and 1500 cm^{-1}. Figure 11 shows clear differences for the patterns for AAA1 and AAA2 which are correctly captured by the NN.

In a conventional force field all these frequencies would be nearly overlapping as the force field parameters for a CO bond do usually not depend on whether a hydrogen is bonded to it or not. In order to capture such an effect, the force field parameters for the CO bond would need to depend on the bonding pattern of the molecule along the dynamics trajectory. Encoding such detail into a conventional force field is difficult and NN-trained PESs offer a natural way to do so.

Another benefit yet to be explored that NN-trained PESs such as PhysNet offer is the possibility to have fluctuating point charges for a molecule without the need to explicitly parametrize the dependence on the geometry. Modeling such effects within an empirical force field is challenging [184].
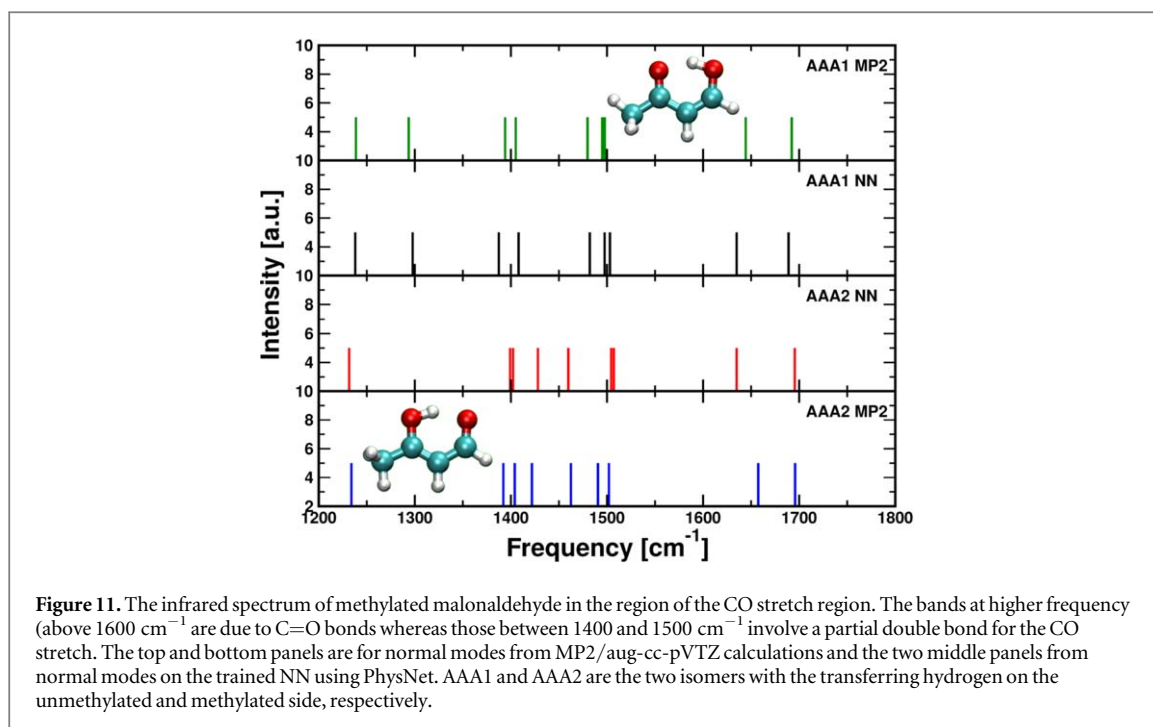
Figure 11. The infrared spectrum of methylated malonaldehyde in the region of the CO stretch region. The bands at higher frequency (above 1600 cm^{-1} are due to C=O bonds whereas those between 1400 and 1500 cm^{-1} involve a partial double bond for the CO stretch. The top and bottom panels are for normal modes from MP2/aug-cc-pVTZ calculations and the two middle panels from normal modes on the trained NN using PhysNet. AAA1 and AAA2 are the two isomers with the transferring hydrogen on the unmethylated and methylated side, respectively.

A final challenge for high-dimensional PESs is including the chemical environment, such as the effect of a solvent. Immersing a chemically reacting system into an environment leads to pronounced changes. As an example, double proton transfer in formic acid dimer (FAD) in the gas phase and in solution is considered. The parametrization used here was adapted to yield the correct infrared spectrum in the gas phase [185]. Recent high-resolution work has confirmed that the barrier of 7.3 kcal mol^{-1} for the gas-phase PES is compatible with the tunneling splitting observed in microwave studies [186]. Such a barrier height makes spontaneous transitions rare. Hence, umbrella sampling simulations were combined with the molecular mechanics with proton transfer (MMPT) force field to determine the free energy barrier for double proton transfer in the gas phase and in solution. As a comparison, the simulations were also carried out by using the Density-Functional Tight-Binding (DFTB) [187, 188] method for the FAD. In both cases the solvent was water represented as the TIP3P model [189].

The free energy barrier in the gas phase is $\Delta G = 5.4$ kcal mol^{-1} which increases to 7.5 kcal mol^{-1} in water, see figure 12. With DFTB3 the barrier height in solution is similar (7.3 kcal mol^{-1}) to that with the MMPT parametrization. In all cases, FAD undergoes a concerted double proton transfer to interconvert between two equivalent forms resulting in a symmetric free energy profile. The nature of the transition state was verified by running 5000 structures from the umbrella sampling simulations at the TS, starting with zero velocity, and propagating them for 1 ps in an *NVE* ensemble. The fraction of reactants and products obtained are 0.54 and 0.46, indicating that the configurations sampled in the umbrella sampling simulations indeed correspond to a transition state and lie midway between reactants and products and are equally likely to relax into either stable state. The solvent distribution is reported in figure 13 and shows that far from the solute it is isotropic (magenta) whereas closer to it is structured (yellow). One of the open question remains to what extent solvent motion contributes to driving double proton transfer; in other words , whether or not double proton transfer and solvent fluctuations are coupled.

From these simulations it is also possible to determine the time to product or reactant which is reported in the inset of figure 12. The most probable time is ∼5 fs with a wide distribution extending out to to 20 fs. This is typical for a waiting time distribution [191] and indicates that multiple degrees of freedom are involved.

The methods discussed in the present work all have their advantages and shortcomings. Depending on the application at hand the methods provide different efficiencies and accuracies and are more or less straightforward to apply. In the following, the three approaches discussed here are compared by looking at them from different perspectives.

- For small gas phase systems such as tri- and tetraatomics, RKHSs, PIPs and NN-based force fields are powerful methods for *accurate* investigations of their reactive dynamics. Empirical force fields are clearly not intended and suitable for this.
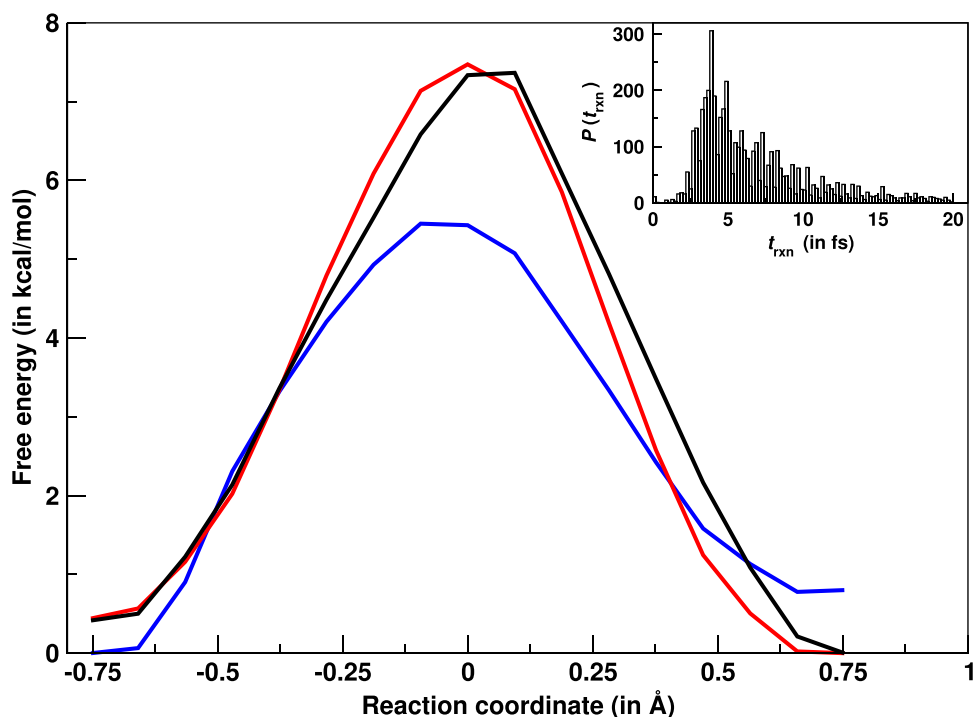
Figure 12. Free energy as a function of reaction coordinate for proton transfer in gaseous and water-solvated FAD. The blue and red curves show the free energy for FAD in the gas and solution phase respectively using the MMPT force field. The energy profile in black is obtained for FAD in solution through DFTB treatment. In all cases, for the umbrella sampling procedure, 17 umbrella windows are located at 0.1 Å intervals and trajectories are propagated for 50 ps. The probability distribution from different umbrellas are recombined using the weighted histogram analysis method (WHAM) [190].
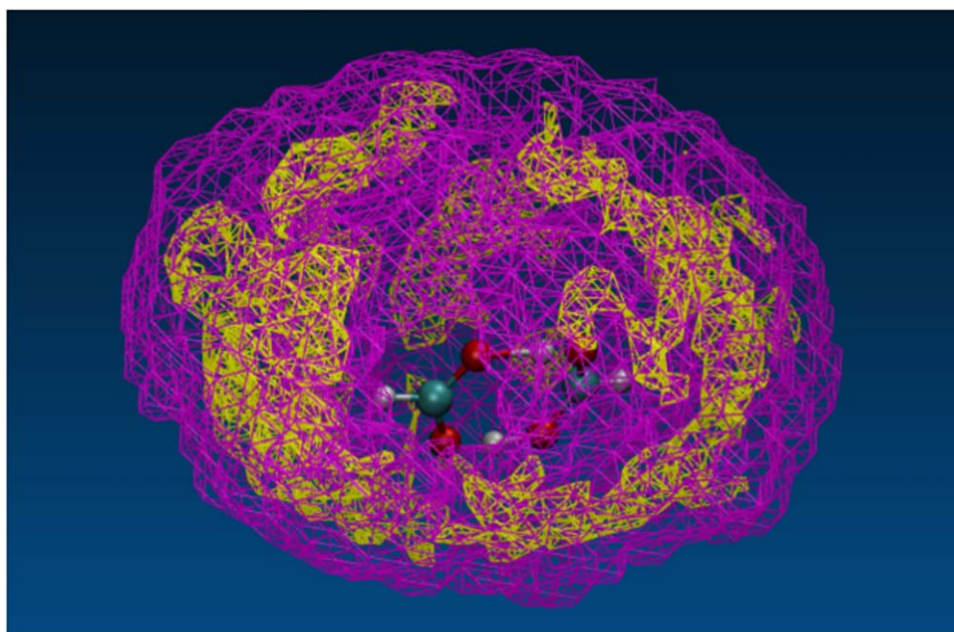
Figure 13. Solvent distribution around FAD for the transition state ensemble from 100 000 transition states sampled from umbrella sampling simulations. The two isosurfaces correspond to fractional occupancies of 10^{-6} (magenta) and 0.01 (yellow). The maximum value of the fractional occupancy at any point is $\sim$0.049.

- For medium-sized molecules (up to $\sim$10 atoms) in the gas phase, reactive MD methods, such as the EVB [49] (not explicitly discussed here or multi state reactive MD), NNs, or suitably parametrized force fields (polarizable or non-polarizable) including multipoles are viable representations. PIPs or RKHSs will eventually become laborious to parametrize and computationally expensive to evaluate.

- Systems with ~10 atoms in solution can be described by refined FFs and reactive MD simulations. NNs, such as Physnet, would be a very attractive possibility, as they include fluctuating charges by construction. Also, capturing changes in the bond character depending on the chemical environment (see discussion of methylated MA above) is readily possible. However, an open technical question is how to include the effect of the environment in training the NN.

- Finally, for macromolecules in solution, such as proteins, either refined reactive FFs or a combination of RKHS and a FF has shown to provide meaningful ways to extend quantitative, reactive simulations to condensed phase systems. Extending such approaches, akin to mixed QM/MM simulations but treating the reactive part with a NN, may provide even better accuracy.

Multidimensional PESs are a powerful way to run high-quality atomistic simulations for gas- and condensed phase systems. Recent progress concerns the accurate, routine representation of PESs based on RKHSs or PIPs. As an exciting alternative, NN-based PESs have also become available. Despite this progress, extension of these techniques to simulations in solution and multiple dimensions remain a challenge. Attractive future possibilities are simulations which capture the changes in local chemistry or in the atomic charges without the need to explicitly parametrize them as a function of geometry. This is possible with approaches as those used in PhysNet.

Acknowledgments

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID iDs

Oliver T Unke ⦿ https://orcid.org/0000-0001-7503-406X
Debasish Koner ⦿ https://orcid.org/0000-0002-5116-4908
Markus Meuwly ⦿ https://orcid.org/0000-0001-7930-8806

References

[1] El Hage K *et al* 2017 Implications of short time scale dynamics on long time processes *Struct. Dyn.* **4** 061507
[2] Minitti M P *et al* 2015 Imaging molecular motion: femtosecond x-ray scattering of an electrocyclic chemical reaction *Phys. Rev. Lett.* **114** 255501
[3] Cui Q and Karplus M 2008 Allostery and cooperativity revisited *Prot. Sci.* **17** 1295–307
[4] Nussinov R and Tsai C-J 2015 Allostery without a conformational change? Revisiting the paradigm *Curr. Opin. Struct. Biol.* **30** 17–24
[5] Cooper A and Dryden D T F 1984 Allostery without conformational change *Eur. Biophys. J.* **11** 103–9
[6] Alon U 2007 *An Introduction to Systems Biology: Design Principles and Biological Circuits* (London: Chapman and Hall/CRC)
[7] Conde M M, Gonzalez M A, Abascal J L F and Vega C 2013 Determining the phase diagram of water from direct coexistence simulations: the phase diagram of the TIP4P/2005 model revisited *J. Chem. Phys.* **139** 154505
[8] Cisneros G A, Wikfeldt K T, Ojamaee L, Lu J, Xu Y, Torabifard H, Bartok A P, Csanyi G, Molinero V and Paesani F 2016 Modeling molecular interactions in water: from pairwise to many body potential energy functions *Chem. Rev.* **116** 7501–28
[9] Xu Z, Shi X, Li J, Lu S and Wang L 2009 Artificial neural network method to construct potential energy surfaces for transition metal nanoparticles: Pt, Au, and Ag *2009 5th Int. Conf. on Natural Computation* pp 86–90
[10] Chiriki S, Jindal S and Bulusu S S 2017 Neural network potentials for dynamics and thermodynamics of gold nanoparticles *J. Chem. Phys.* **146** 084314
[11] Jonayat A S M, van Duin A C T and Janik M J 2018 Discovery of descriptors for stable monolayer oxide coatings through Machine Learning *ACS Appl. Energy Mater.* **1** 6217–26
[12] Pilania G, Gubernatis J and Lookman T 2017 Multi-fidelity machine learning models for accurate bandgap predictions of solids *Comput. Mater. Sci.* **129** 156–63
[13] Rajan A C, Mishra A, Satsangi S, Vaish R, Mizuseki H, Lee K-R and Singh A K 2018 Machine-learning-assisted accurate band gap predictions of functionalized MXene *Chem. Mater.* **30** 4031–8
[14] Davies D W, Butler K T, Skelton J M, Xie C, Oganov A R and Walsh A 2018 Computer-aided design of metal chalcohalide semiconductors: from chemical composition to crystal structure *Chem. Sci.* **9** 1022–30
[15] Jinnouchi R, Karsai F and Kresse G 2019 On-the-fly machine learning force field generation: Application to melting points *Phys. Rev. B* **100** 014105
[16] Cui Q 2016 Perspective: quantum mechanical methods in biochemistry and biophysics *J. Chem. Phys.* **145** 140901

[17] Senn H M and Thiel W 2009 QM/MM methods for biomolecular systems *Angew. Chem. Int. Ed.* **48** 1198–229

[18] Roston D, Demapan D and Cui Q 2016 Leaving group ability observably affects transition state structure in a single enzyme active site *J. Am. Chem. Soc.* **138** 7386–94

[19] Kulik H J, Zhang J, Klinman J P and Martinez T J 2016 How large should the QM region Be in QM/MM calculations? The case of catechol O-methyltransferase *J. Phys. Chem.* B **120** 11381–94

[20] Feig M, Yu I, Wang P-H, Nawrocki G and Sugita Y 2017 Crowding in cellular environments at an atomistic level from computer simulations *J. Phys. Chem.* B **121** 8009–25

[21] Rasmussen T D, Ren P, Ponder J W and Jensen F 2007 Force field modeling of conformational energies: importance of multipole moments and intramolecular polarization *Int. J. Quant. Chem.* **107** 1390–5

[22] Kramer C, Gedeck P and Meuwly M 2013 Multipole-based force fields from *ab initio* interaction energies and the need for jointly refitting all intermolecular parameters *J. Chem. Theory Comput.* **9** 1499–511

[23] Bereau T, Kramer C and Meuwly M 2013 Leveraging symmetries of static atomic multipole electrostatics in molecular dynamics simulations *J. Chem. Theory Comput.* **9** 5450–9

[24] Shi Y, Xia Z, Zhang J, Best R, Wu C, Ponder J W and Ren P 2013 Polarizable atomic multipole-based AMOEBA force field for proteins *J. Chem. Theory Comput.* **9** 4046–63

[25] Hédin F, El Hage K and Meuwly M 2016 A toolkit to fit nonbonded parameters from and for condensed phase simulations *J. Chem. Inf. Model.* **56** 1479–89

[26] Lopes P E M, Huang J, Shim J, Luo Y, Li H, Roux B and MacKerell A D Jr. 2013 Polarizable force field for peptides and proteins based on the classical drude oscillator *J. Chem. Theory Comput.* **9** 5430–49

[27] Rappe A K, Casewit C J, Colwell K S, Goddard W A and Skiff W M 1992 UFF, a full periodic-table force-field for molecular mechanics and molecular-dynamics simulations *J. Am. Chem. Soc.* **114** 10024–35

[28] Sun H 1998 COMPASS: an *ab initio* force-field optimized for condensed-phase applications—overview with details on alkane and benzene compounds *J. Phys. Chem.* B **102** 7338–64

[29] Mayo S L, Olafson B D and Goddard W A 1990 DREIDING—a generic force-field for molecular simulations *J. Phys. Chem.* **94** 8897–909

[30] Landis C R, Cleveland T and Firman T K 1998 Valence bond concepts applied to the molecular mechanics description of molecular shapes: III. Applications to transition metal alkyls and hydrides *J. Am. Chem. Soc.* **120** 2641–9

[31] Firman T K and Landis C R 2001 Valence bond concepts applied to the molecular mechanics description of molecular shapes: IV. Transition metals with pi-bonds *J. Am. Chem. Soc.* **123** 11728–42

[32] Tubert-Brohman I, Schmid M and Meuwly M 2009 A molecular mechanics force field for octahedral organometallic compounds with inclusion of the trans influence *J. Chem. Theory Comput.* **5** 530

[33] Schmid M H, Das A K, Landis C R and Meuwly M 2018 Multi-state VALBOND for atomistic simulations of hypervalent molecules, metal complexes, and reactions *J. Chem. Theory Comput.* **14** 3565–78

[34] Braams B J and Bowman J M 2009 Permutationally invariant potential energy surfaces in high dimensionality *Int. Rev. Phys. Chem.* **28** 577–606

[35] Xie Z and Bowman J M 2010 Permutationally invariant polynomial basis for molecular energy surface fitting via monomial symmetrization *J. Chem. Theory Comput.* **6** 26–34

[36] Jin Z, Braams B and Bowman J 2006 An *ab initio* based global potential energy surface describing $CH_5^+ \rightarrow CH_3^+ + H^2$ *J. Phys. Chem.* A **110** 1569–74

[37] Huang X, Johnson L, Bowman J and McCoy A 2006 Deuteration effects on the structure and infrared spectrum of CH_5 *J. Am. Chem. Soc.* **128** 3478–9

[38] Huang X, Braams B J, Bowman J M, Kelly R E A, Tennyson J, Groenenboom G C and van der Avoird A 2008 New *ab initio* potential energy surface and the vibration-rotation-tunneling levels of $(H_2O)_2$ and $(D_2O)_2$ *J. Chem. Phys.* **128** 034312

[39] Wang Y, Huang X, Shepler B C, Braams B J and Bowman J M 2011 Flexible, *ab initio* potential, and dipole moment surfaces for water: I. Tests and applications for clusters up to the 22-mer *J. Chem. Phys.* **134** 094509

[40] Shepler B C, Braams B J and Bowman J M 2007 Quasiclassical trajectory calculations of acetaldehyde dissociation on a global potential energy surface indicate significant non-transition state dynamics *J. Phys. Chem.* A **111** 8282–5

[41] Varandas A J C 1988 Intermolecular and intramolecular potential- topographical aspects, calculation, and functional representation via a double many-body expansion method *Adv. Chem. Phys.* **74** 255–338

[42] Goncalves C E M, Galvao B R L, Mota V C, Braga J P and Varandas A J C 2018 Accurate explicit-correlation-MRCI-based DMBE potential-energy surface for ground-state CNO *J. Phys. Chem.* A **122** 4198–207

[43] Koner D, Bemish R J and Meuwly M 2018 The $C(^3P) + NO(X^2\Pi) \rightarrow O(^3P) + CN(X\ ^2\Sigma^+)$, $N(^2D)/N(^4S) + CO(X\ ^1\Sigma^+)$ reaction: rates, branching ratios, and final states from 15 K to 20 000 K *J. Chem. Phys.* **149** 094305

[44] Collins M A 2002 Molecular potential-energy surfaces for chemical reaction dynamics *Theor. Chem. Acc.* **108** 313–24

[45] Dawes R, Passalacqua A, Wagner A F, Sewell T D, Minkoff M and Thompson D L 2009 Interpolating moving least-squares methods for fitting potential energy surfaces: using classical trajectories to explore configuration space *J. Chem. Phys.* **130** 144107

[46] Quintas-Sanchez E and Dawes R 2019 AUTOSURF: a freely available program to construct potential energy surfaces *J. Chem. Inf. Model.* **59** 262–71

[47] Nagy T, Yosa Reyes J and Meuwly M 2014 Multisurface adiabatic reactive molecular dynamics *J. Chem. Theory Comput.* **10** 1366–75

[48] Danielsson J and Meuwly M 2008 Atomistic simulation of adiabatic reactive processes based on multi-state potential energy surfaces *J. Chem. Theory Comput.* **4** 1083

[49] Warshel A and Weiss R 1980 An empirical valence bond approach for comparing reactions in solutions and in enzymes *J. Am. Chem. Soc.* **102** 6218–26

[50] Rupp M, Tkatchenko A, Müller K-R and Von Lilienfeld O A 2012 Fast and accurate modeling of molecular atomization energies with machine learning *Phys. Rev. Lett.* **108** 058301

[51] Montavon G, Rupp M, Gobre V, Vazquez-Mayagoitia A, Hansen K, Tkatchenko A, Müller K-R and Von Lilienfeld O A 2013 Machine learning of molecular electronic properties in chemical compound space *New J. Phys.* **15** 095003

[52] Hansen K, Montavon G, Biegler F, Fazli S, Rupp M, Scheffler M, Von Lilienfeld O A, Tkatchenko A and Müller K-R 2013 Assessment and validation of machine learning methods for predicting molecular atomization energies *J. Chem. Theory Comput.* **9** 3404–19

[53] Hansen K, Biegler F, Ramakrishnan R, Pronobis W, Von Lilienfeld O A, Müller K-R and Tkatchenko A 2015 Machine learning predictions of molecular properties: accurate many-body potentials and nonlocality in chemical space *J. Phys. Chem. Lett.* **6** 2326–31

[54] Samuel A L 2000 Some studies in machine learning using the game of checkers *IBM J. Res. Dev.* **44** 206–26

[55] Schölkopf B, Smola A and Müller K-R 1997 Kernel principal component analysis *Int. Conf. on Artificial Neural Networks* (Berlin: Springer) pp 583–8

[56] Boser B E, Guyon I M and Vapnik V N 1992 A training algorithm for optimal margin classifiers *Proc. 5th Annual Workshop on Computational Learning Theory* pp 144–52

[57] Schölkopf B, Smola A and Müller K-R 1998 Nonlinear component analysis as a kernel eigenvalue problem *Neural Comput.* **10** 1299–319

[58] Theodoridis S and Koutroumbas K E A 2008 *Pattern Recognition* (Amsterdam: Elsevier)

[59] Behler J 2011 Atom-centered symmetry functions for constructing high-dimensional neural network potentials *J. Chem. Phys.* **134** 074106

[60] Bartók A P, Kondor R and Csányi G 2013 On representing chemical environments *Phys. Rev.* B **87** 184115

[61] Faber F A, Christensen A S, Huang B and von Lilienfeld O A 2018 Alchemical and structural distribution based representation for universal quantum machine learning *J. Chem. Phys.* **148** 241717

[62] Christensen A S, Bratholm L A, Faber F A, Glowacki D R and von Lilienfeld O A 2019 FCHL revisited: faster and more accurate quantum machine learning arXiv:1909.01946

[63] Schütt K, Glawe H, Brockherde F, Sanna A, Müller K and Gross E 2014 How to represent crystal structures for machine learning: towards fast prediction of electronic properties *Phys. Rev.* B **89** 205118

[64] Faber F, Lindmaa A, von Lilienfeld O A and Armiento R 2015 Crystal structure representations for machine learning models of formation energies *Int. J. Quantum Chem.* **115** 1094–101

[65] Faber F A, Lindmaa A, Von Lilienfeld O A and Armiento R 2016 Machine learning energies of 2 million elpasolite (ABC_2D_6) crystals *Phys. Rev. Lett.* **117** 135502

[66] Schölkopf B, Herbrich R and Smola A J 2001 A generalized representer theorem *Int. Conf. on Computational Learning Theory* (Berlin: Springer) pp 416–26

[67] Berlinet A and Thomas-Agnan C 2011 *Reproducing Kernel Hilbert Spaces in Probability and Statistics* (Dordrecht: Springer Science & Business Media)

[68] Hollebeek T, Ho T-S and Rabitz H 1999 Constructing multidimensional molecular potential energy surfaces from *ab initio* data *Ann. Rev. Phys. Chem.* **50** 537–70

[69] Soldan P and Hutson J 2000 On the long-range and short-range behavior of potentials from reproducing kernel Hilbert space interpolation *J. Chem. Phys.* **112** 4415–6

[70] Müller K-R, Mika S, Rätsch G, Tsuda K and Schölkopf B 2001 An introduction to kernel-based learning algorithms *IEEE Trans. Neural Netw.* **12** 181–201

[71] Hofmann T, Schölkopf B and Smola A J 2008 Kernel methods in machine learning *Ann. Stat.* **2008** 1171–220

[72] Golub G H and Van Loan C F 2012 *Matrix Computations* vol 3 (Baltimore: JHU Press)

[73] Tikhonov A N, Arsenin V I and John F 1977 *Solutions of Ill-Posed Problems* vol 14 (Washington, DC: Winston)

[74] Rasmussen C and Williams C 2005 *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)* (Berlin: Springer) pp 1–247

[75] McCulloch W S and Pitts W 1943 A logical calculus of the ideas immanent in nervous activity *Bull. Math. Biophys.* **5** 115–33

[76] Kohonen T 1988 An introduction to neural computing *Neural Netw.* **1** 3–16

[77] Abdi H 1994 A neural network primer *J. Biol. Syst.* **2** 247–81

[78] Bishop C M 1995 *Neural Networks for Pattern Recognition* (Oxford: Oxford University Press)

[79] Clark J W 1999 *Scientific Applications of Neural Nets* (Berlin: Springer) pp 1–96

[80] Ripley B D 2007 *Pattern Recognition and Neural Networks* (Cambridge: Cambridge University Press)

[81] Haykin S S 2009 *Neural Networks and Learning Machines* vol 3 (Upper Saddle River, NJ: Pearson)

[82] Cybenko G 1989 Approximation by superposition of sigmoidal functions *Math. Control Signals Syst.* **2** 303–14

[83] Hornik K 1991 Approximation capabilities of multilayer feedforward networks *Neural Netw.* **4** 251–7

[84] Eldan R and Shamir O 2016 The power of depth for feedforward neural networks *Conf. on Learning Theory* pp 907–40

[85] Blank T B, Brown S D, Calhoun A W and Doren D J 1995 Neural network models of potential energy surfaces *J. Chem. Phys.* **103** 4129–37

[86] Brown D F, Gibbs M N and Clary D C 1996 Combining *ab initio* computations, neural networks, and diffusion Monte Carlo: an efficient method to treat weakly bound molecules *J. Chem. Phys.* **105** 7597–604

[87] Tafeit E, Estelberger W, Horejsi R, Moeller R, Oettl K, Vrecko K and Reibnegger G 1996 Neural networks as a tool for compact representation of *ab initio* molecular potential energy surfaces *J. Mol. Graph.* **14** 12–8

[88] No K T, Chang B H, Kim S Y, Jhon M S and Scheraga H A 1997 Description of the potential energy surface of the water dimer with an artificial neural network *Chem. Phys. Lett.* **271** 152–6

[89] Prudente F V and Neto J S 1998 The fitting of potential energy surfaces using neural networks. Application to the study of the photodissociation processes *Chem. Phys. Lett.* **287** 585–9

[90] Manzhos S and Carrington T Jr 2006 A random-sampling high dimensional model representation neural network for building potential energy surfaces *J. Chem. Phys.* **125** 084109

[91] Manzhos S and Carrington T Jr 2007 Using redundant coordinates to represent potential energy surfaces with lower-dimensional functions *J. Chem. Phys.* **127** 014103

[92] Malshe M, Narulkar R, Raff L, Hagan M, Bukkapatnam S, Agrawal P and Komanduri R 2009 Development of generalized potential-energy surfaces using many-body expansions, neural networks, and moiety energy approximations *J. Chem. Phys.* **130** 184102

[93] Behler J and Parrinello M 2007 Generalized neural-network representation of high-dimensional potential-energy surfaces *Phys. Rev. Lett.* **98** 146401

[94] Khorshidi A and Peterson A A 2016 Amp: a modular approach to machine learning in atomistic simulations *Comput. Phys. Commun.* **207** 310–24

[95] Artrith N, Urban A and Ceder G 2017 Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species *Phys. Rev.* B **96** 014112

[96] Unke O T and Meuwly M 2018 A reactive, scalable, and transferable model for molecular energies from a neural network approach based on local information *J. Chem. Phys.* **148** 241708

[97] Smith J S, Isayev O and Roitberg A E 2017 ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost *Chem. Sci.* **8** 3192–203

[98] Yao K, Herr J E, Toth D W, Mckintyre R and Parkhill J 2018 The TensorMol-0.1 model chemistry: a neural network augmented with long-range physics *Chem. Sci.* **9** 2261–9

[99] Gilmer J, Schoenholz S S, Riley P F, Vinyals O and Dahl G E 2017 Neural message passing for quantum chemistry *Int. Conf. on Machine Learning* pp 1263–72

[100] Schütt K T, Arbabzadah F, Chmiela S, Müller K R and Tkatchenko A 2017 Quantum-chemical insights from deep tensor neural networks *Nat. Commun.* **8** 13890

[101] Schuett K T, Kindermans P-J, Felix H E S, Chmiela S, Tkatchenko A and Mueller K-R 2017 A continuous-filter convolutional neural network for modeling quantum interactions *Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017; 31st Annual Conf. on Neural Information Processing Systems (NIPS) (Long Beach, CA)* 991–1001

[102] Lubbers N, Smith J S and Barros K 2018 Hierarchical modeling of molecular energies using a deep neural network *J. Chem. Phys.* **148** 241715

[103] Unke O T and Meuwly M 2019 PhysNet: a neural network for predicting energies, forces, dipole moments and partial charges *J. Chem. Theory Comput.* **15** 3678–93

[104] Schütt K T, Gastegger M, Tkatchenko A and Müller K-R 2019 *Quantum-Chemical Insights from Interpretable Atomistic Neural Networks* (Cham, Switzerland: Springer) pp 311–30

[105] Armenise I and Esposito F 2015 N_2, O_2, NO state-to-state vibrational kinetics in hypersonic boundary layers: The problem of rescaling rate coefficients to uniform vibrational ladders *Chem. Phys.* **446** 30–46

[106] Dean A J, Hanson R K and Bowman C T 1991 A shock tube study of reactions of carbon atoms and methylidyne with nitric oxide including product channel measurements *J. Phys. Chem.* **95** 3180

[107] Bergeat A, Calvo T, Dorthe G and Loison J C 1999 Fast-flow study of the C+NO and C+O_2 reactions *Chem. Phys. Lett.* **308** 7

[108] Braun W, Bass A M, Davis D D and Simmons J D 1969 Flash photolysis of carbon suboxide: absolute rate constants for reactions of C(3P) and C(1D) with H_2, N_2, CO, NO, O_2 and CH_4 *Proc. R. Soc. A* **312** 417–34

[109] Landau L D 1932 On the theory of transfer of energy at collisions II *Phys. Z.* **2** 46

[110] Zener C 1932 Non-adiabatic crossing of energy levels *Proc. R. Soc. A* **137** 696

[111] Belyaev A K and Lebedev O V 2011 Nonadiabatic nuclear dynamics of atomic collisions based on branching classical trajectories *Phys. Rev. A* **84** 014701

[112] Belyaev A K, Lasser C and Trigila G J 2014 Landau–Zener type surface hopping algorithms *Chem. Phys.* **140** 224108

[113] Bird G A 1994 *Molecular Gas Dynamics and the Direct Simulation of Gas Flows* (Oxford: Clarendon)

[114] Abrahamsson E, Andersson S, Marković N and Nyman G 2008 An improved potential energy surface for the C + NO reaction *Phys. Chem. Chem. Phys.* **10** 4400–9

[115] Husain D and Kirsch L J 1971 Kinetic investigation of C(2^3P_J) by photoelectric measurement of the attenuation of atomic emission in the vacuum ultra-violet *Chem. Phys. Lett.* **8** 543–6

[116] Husain D and Young A N 1975 Kinetic investigation of ground state carbon atoms, C(2^3P_J) *J. Chem. Soc. Faraday Trans.* II **71** 525–31

[117] Chastaing D, Le Picard S D and Sims I R 2000 Direct kinetic measurements on reactions of atomic carbon, C(3P),C(3P), with O_2 and NO at temperatures down to 15 K *J. Chem. Phys.* **112** 8466

[118] Rivero U, Unke O T, Meuwly M and Willitsch S 2019 A computational study of the Diels-Alder reactions between 2,3-dibromo-1,3-butadiene and maleic anhydride *J. Chem. Phys.* **151** 104301

[119] Chang Y-P, Dlugolecki K, Küpper J, Rösch D, Wild D and Willitsch S 2013 *Science* **342** 98

[120] Willitsch S 2017 Chemistry with controlled ions *Adv. Chem. Phys.* **162** 307–40

[121] de Souza M A F, Ventura E, do Monte S A, Riveros J M and Longo R L J 2016 *Comput. Chem.* **37** 701

[122] Black K, Liu P, Xu L, Doubleday C and Houk K N 2012 *Proc. Natl Acad. Sci.* **109** 12860

[123] Tan J S J, Hirvonen V and Paton R S 2018 *Org. Lett.* **20** 2821

[124] Wang Z, Hirschi J S and Singleton D A 2009 *Angew. Chem. Int. Ed.* **48** 9156

[125] Liu F, Yang Z, Mei Y and Houk K N 2016 *J. Phys. Chem. B* **120** 6250

[126] Soto-Delgado J, Tapia R A and Torras J 2016 *J. Chem. Theory Comput.* **12** 4735

[127] Unke O T, Brickel S and Meuwly M 2019 Sampling reactive regions in phase space by following the minimum dynamic path *J. Chem. Phys.* **150** 074107

[128] Brickel S, Das A K, Unke O T, Turan H T and Meuwly M 2019 Reactive molecular dynamics for the [Cl-CH_3-Br]$^-$ reaction in the gas phase and in solution: a comparative study using empirical and neural network force fields *Electron. Struct.* **1** 024002

[129] Soloviov M, Das A K and Meuwly M 2016 Structural interpretation of metastable states in myoglobin-NO *Angew. Chem. Intern. Ed.* **55** 10126–30

[130] Unke O T and Meuwly M 2017 Toolkit for the construction of reproducing kernel-based representations of data: application to multidimensional potential energy surfaces *J. Chem. Inf. Model.* **57** 1923–31

[131] Cornelius P, Hochstrasser R and Steele A 1983 Ultrafast relaxatio nin picosecond photolysis of nitrosylhemoglobin *J. Mol. Biol.* **163** 119–28

[132] Petrich J W, Lambry J C, Kuczera K, Karplus M, Poyart C and Martin J L 1991 Ligand binding and protein relaxation in heme proteins: a room temperature analysis of nitric oxide geminate recombination *Biochemistry* **30** 3975–87

[133] Ionascu D, Gruia F, Ye X, Yu A, Rosca F, Demidov C B A, Olson J S and Champion P M 2005 Temperature-dependent studies of NO recombination to heme and heme proteins *J. Am. Chem. Soc.* **127** 16921–34

[134] Kruglik S G, Yoo B-K, Franzen S, Vos M H, Martin J-L and Negrerie M 2010 Picosecond primary structural transition of the heme is retarded after nitric oxide binding to heme proteins *Proc. Natl Acad. Sci.* **107** 13678–83

[135] Kim J, Park J, Lee T and Lim M 2012 Dynamics of geminate rebinding of NO with cytochrome c in aqueous solution using femtosecond vibrational spectroscopy *J. Phys. Chem. B* **116** 13663–71

[136] Yoo B-K, Kruglik S G, Lamarre I, Martin J-L and Negrerie M 2012 Absorption band III kinetics probe the picosecond heme iron motion triggered by nitric oxide binding to hemoglobin and myoglobin *J. Phys. Chem. B* **116** 4106–14

[137] Silatani M, Lima F A, Penfold T J, Rittmann J, Reinhard M E, Rittmann-Frank H M, Borca C, Grolimund D, Milne C J and Chergui M 2015 NO binding kinetics in myoglobin investigated by picosecond Fe K-edge absorption spectroscopy *Proc. Natl Acad. Sci.* **112** 12922–7

[138] Kim S, Jin G and Lim M 2004 Dynamics of geminate recombination of no with myoglobin in aqueous solution probed by femtosecond mid-IR spectroscopy *J. Phys. Chem. B* **108** 20366–75

[139] Meuwly M, Becker O M, Stote R and Karplus M 2002 NO rebinding to myoglobin: a reactive molecular dynamics study *Biophys. Chem.* **98** 183–207

[140] Ye X, Demidov A and Champion P 2002 Measurements of the photodissociation quantum yields of MbNO and MbO_2 and the vibrational relaxation of the six-coordinate heme species *J. Am. Chem. Soc.* **124** 5914–24

[141] Nutt D R and Meuwly M 2006 Studying reactive processes with classical dynamics: Rebinding dynamics in MbNO *Biophys. J.* **90** 1191–201

[142] Merchant K A, Noid W G, Thompson D E, Akiyama R, Loring R F and Fayer M D 2003 Structural assignments and dynamics of the A substates of MbCO: spectrally resolved vibrational echo experiments and molecular dynamics simulations *J. Phys. Chem.* B **107** 4–7

[143] Claisen L 1912 Über Umlagerung von Phenol-allyläthern in C-Allyl-phenole *Chem. Ber.* **45** 3157–66

[144] Iwakura I 2011 The experimental visualisation of molecular structural changes during both photochemical and thermal reactions by real-time vibrational spectroscopy *Phys. Chem. Chem. Phys.* **13** 5546–55

[145] Coates R M, Rogers B D, Hobbs S J, Curran D P and Peck D R 1987 Synthesis and claisen rearrangement of alkoxyallyl enol ethers. Evidence for a dipolar transition state *J. Am. Chem. Soc.* **35** 2601–5

[146] Ziegler F E 1988 The thermal, aliphatic claisen rearrangement *Chem. Rev.* **88** 1423–52

[147] Severance D L and Jorgensen W L 1992 Effects of hydration on the claisen rearrangement of allyl vinyl ether from computer simulations *J. Am. Chem. Soc.* **114** 10966–8

[148] Guest J M, Craw J S, Vincent M A and Hillier I H 1997 The effect of water on the claisen rearrangement of allyl vinyl ether: theoretical methods including explicit solvent and electron correlation *Perkin Trans.* 2 71–4

[149] Cramer C J and Truhlar D G 1992 ChemInform abstract: what causes aqueous acceleration of the claisen rearrangement? *J. Am. Chem. Soc.* **114** 8794–9

[150] Kast P, Asif-Ullah M and Hilvert D 1996 Is chorismate mutase a prototypic entropy trap? - activation parameters for the bacillus subtilis enzyme *Tetrahedron Lett.* **37** 2691–4

[151] Ranaghan K E, Ridder L, Szefczyk B, Sokalski W A, Hermann J C and Mulholland A J 2003 Insights into enzyme catalysis from QM/MM modelling: transition state stabilization in chorismate mutase *Mol. Phys.* **101** 2695–714

[152] Lever G, Cole D J, Lonsdale R, Ranaghan K E, Wales D J, Mulholland A J, Skylaris C K and Payne M C 2014 Large-scale density functional theory transition state searching in enzymes *J. Phys. Chem. Lett.* **5** 3614–9

[153] Martí S, Andrés J, Moliner V, Silla E, Tuñón I and Bertrán J 2010 Theoretical QM/MM studies of enzymatic pericyclic reactions *Interdiscip. Sci. Comput. Life Sci.* **2** 115–31

[154] Ferrer S, Martí S, Andrés J, Moliner V, Tuñón I and Bertrán J 2011 Molecular mechanism of chorismate mutase activity of promiscuos MbtI *Theor. Chem. Acc.* **128** 601–7

[155] Martí S, Andrés J, Moliner V, Silla E, Tuñón I, Bertrán J and Field M J 2001 A hybrid potential reaction path and free energy study of the chorismate mutase reaction *J. Am. Chem. Soc.* **123** 1709–12

[156] Roca M, Vardi-Kilshtain A and Warshel A 2009 Toward accurate screening in computer-aided enzyme design *Biochemistry* **48** 3046–56

[157] Madurga S and Vilaseca E 2001 SCRF study of the conformational equilibrium of chorismate in water *Phys. Chem. Chem. Phys.* **3** 3548–54

[158] Davidson M M, Guest J M, Craw J S, Hillier I H and Vincent M A 1997 Conformational and solvation aspects of the chorismate—prephenate rearrangement studied by *ab initio* electronic structure and simulation methods *Perkin Trans.* 2 1395–400

[159] Wiest O and Houk K N 1995 Stabilization of the transition state of the chorismate-prephenate rearrangement: an *ab initio* study of enzyme and antibody catalysis *J. Am. Chem. Soc.* **117** 11628–39

[160] Hur S and Bruice T C 2002 The mechanism of catalysis of the chorismate to prephenate reaction by the escherichia coli mutase enzyme *Proc. Natl Acad. Sci. USA* **99** 1176–81

[161] Vance R and Rondan N 1988 Transition structures for the claisen rearrangement *J. Am. Chem. Soc.* **110** 2314–5

[162] Claeyssens F, Ranaghan K E, Lawan N, Macrae S J, Manby F R, Harvey J N and Mulholland A J 2011 Analysis of chorismate mutase catalysis by QM/MM modelling of enzyme-catalysed and uncatalysed reactions *Org. Biomol. Chem.* **9** 1578–90

[163] Ranaghan K E, Ridder L, Szefczyk B, Sokalski W A, Hermann J C and Mulholland A J 2004 Transition state stabilization and substrate strain in enzyme catalysis: *ab initio* QM/MM modelling of the chorismate mutase reaction *Org. Biomol. Chem.* **2** 968–80

[164] Carlson H A and Jorgensen W L 1996 Monte Carlo investigations of solvent effects on the chorismate to prephenate rearrangement *J. Am. Chem. Soc.* **118** 8475–84

[165] Andrews P R, Smith G D and Young I G 1973 Transition-state stabilization and enzymic catalysis. Kinetic and molecular orbital studies of the rearrangement of chorismate to prephenate *Biochemistry* **12** 3492–8

[166] Brickel S and Meuwly M 2019 Molecular determinants for rate acceleration in the claisen rearrangement reaction *J. Phys. Chem.* B **123** 448–56

[167] Wright S, DeClue M, Mandal A, Lee L, Wiest O, Cleland W and Hilvert D 2005 Isotope effects on the enzymatic and nonenzymatic reactions of chorismate *J. Am. Chem. Soc.* **127** 12957–64

[168] Addadi L, Jaffe E K and Knowles J R 1983 Secondary tritium isotope effects as probes of the enzymic and nonenzymic conversion of chorismate to prephenate *Biochemistry* **22** 4494–501

[169] O'Ferrall R A M 1970 Relationships between E2 and E1cB mechanisms of β-elimination *J. Chem. Soc.* # 274–7

[170] Jencks W P 1972 General acid-base catalysis of complex reactions in water *Chem. Rev.* **72** 705–18

[171] Ruddigkeit L, Van Deursen R, Blum L C and Reymond J-L 2012 Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17 *J. Chem. Inf. Model.* **52** 2864–75

[172] Faber F A, Hutchison L, Huang B, Gilmer J, Schoenholz S S, Dahl G E, Vinyals O, Kearnes S, Riley P F and Von Lilienfeld O A 2017 Prediction errors of molecular machine learning models lower than hybrid DFT error *J. Chem. Theory Comput.* **13** 5255–64

[173] Ramakrishnan R, Dral P O, Rupp M and Von Lilienfeld O A 2014 Quantum chemistry structures and properties of 134 kilo molecules *Sci. Data* **1** 140022

[174] Schütt K, Gastegger M, Tkatchenko A, Müller K-R and Maurer R 2019 Unifying machine learning and quantum chemistry-a deep neural network for molecular wavefunctions *Nat. Commun.* **10** 1–10

[175] Christensen A S, Faber F A and von Lilienfeld O A 2019 Operators in quantum machine learning: response properties in chemical space *J. Chem. Phys.* **150** 064105

[176] Paukku Y, Yang K R, Varga Z and Truhlar D G 2013 Global *ab initio* ground-state potential energy surface of N_4 *J. Chem. Phys.* **139** 044309

[177] Paukku Y, Yang K R, Varga Z and Truhlar D G 2014 Erratum: global *ab initio* ground-state potential energy surface of N_4 *J. Chem. Phys.* **140** 019903

[178] Qu C and Bowman J M 2019 A fragmented, permutationally invariant polynomial approach for potential energy surfaces of large molecules: application to N-methyl acetamide *J. Chem. Phys.* **150** 141101

[179] Chmiela S, Tkatchenko A, Sauceda H E, Poltavsky I, Schütt K T and Müller K-R 2017 Machine learning of accurate energy-conserving molecular force fields *Sci. Adv.* **3** e1603015

[180]   Chmiela S, Sauceda H E, Poltavsky I, Müller K-R and Tkatchenko A 2019 sGDML: Constructing accurate and data efficient molecular force fields using machine learning *Comput. Phys. Commun.* **240** 38–45

[181]   Sauceda H E, Chmiela S, Poltavsky I, Müller K-R and Tkatchenko A 2019 Construction of machine learned force fields with quantum chemical accuracy: applications and chemical insights arXiv:1909.08565

[182]   Vargas-Hernandez R A, Guan Y, Zhang D H and Krems R V 2019 Bayesian optimization for the inverse scattering problem in quantum reaction dynamics *New. J. Phys.* **21** 022001

[183]   Xu Z-H and Meuwly M 2017 Vibrational spectroscopy and proton transfer dynamics in protonated oxalate *J. Phys. Chem.* A **121** 5389–98

[184]   Patel S and Brooks C 2004 CHARMM fluctuating charge force field for proteins: I. Parameterization and application to bulk organic liquid simulations *J. Comput. Chem.* **25** 1–15

[185]   Mackeprang K, Xu Z-H, Maroun Z, Meuwly M and Kjaergaard H G 2016 Spectroscopy and dynamics of double proton transfer in formic acid dimer *Phys. Chem. Chem. Phys.* **18** 24654–62

[186]   Li W, Evangelisti L, Gou Q, Caminati W and Meyer R 2019 The barrier to proton transfer in the dimer of formic acid: a pure rotational study *Angew. Chem. Int. Ed.* **58** 859–65

[187]   Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S and Seifert G 1998 Self-consistent-charge density-functional tight-binding method for simulations of complex materials proper ties *Phys. Rev.* B **58** 7260–8

[188]   Gaus M, Cui Q and Elstner M 2014 Density functional tight binding: application to organic and biological molecules *WIREs: Comput. Mol. Sci.* **4** 49–61

[189]   Jorgensen W L, Chandrasekhar J, Madura J D, Impey R W and Klein M L 1983 Comparison of simple potential functions for simulating liquid water *J. Chem. Phys.* **79** 926–35

[190]   Kumar S, Bouzida D, Swendsen R H, Kollman P A and Rosenberg J M 1992 The weighted histogram analysis method for free-energy calculations on biomolecules: I. The method *J. Comput. Chem.* **13** 1011–21

[191]   Nadarajah S 2007 The waiting time distribution *Comput. Ind. Eng.* **53** 693–9