Article

Development of Range-Corrected Deep Learning Potentials for Fast, Accurate Quantum Mechanical/Molecular Mechanical Simulations of Chemical Reactions in Solution

Jinzhe Zeng, Timothy J. Giese, Şölen Ekesan, and Darrin M. York*

Cite This: J. Che	m. Theory Comput. 2021, 17, 69	993–7009	Read Online	
ACCESS	III Metrics & More	Ar	ticle Recommendations	s Supporting Information

ABSTRACT: We develop a new deep potential—range correction (DPRc) machine learning potential for combined quantum mechanical/molecular mechanical (QM/MM) simulations of chemical reactions in the condensed phase. The new range correction enables short-ranged QM/MM interactions to be tuned for higher accuracy, and the correction smoothly vanishes within a specified cutoff. We further develop an active learning procedure for robust neural network training. We test the DPRc model and training procedure against a series of six nonenzymatic phosphoryl transfer reactions in solution that are important in mechanistic studies of RNA-cleaving enzymes. Specifically, we apply DPRc corrections to a base QM model and



test its ability to reproduce free-energy profiles generated from a target QM model. We perform these comparisons using the MNDO/d and DFTB2 semiempirical models because they differ in the way they treat orbital orthogonalization and electrostatics and produce free-energy profiles which differ significantly from each other, thereby providing us a rigorous stress test for the DPRc model and training procedure. The comparisons show that accurate reproduction of the free-energy profiles requires correction of the QM/MM interactions out to 6 Å. We further find that the model's initial training benefits from generating data from temperature replica exchange simulations and including high-temperature configurations into the fitting procedure, so the resulting models are trained to properly avoid high-energy profiles made from the target QM/MM simulations. The DPRc model was further demonstrated to be transferable to 2D free-energy surfaces and 1D free-energy profiles that were not explicitly considered in the training. Examination of the computational performance of the DPRc model showed that it was fairly slow when run on CPUs but was sped up almost 100-fold when using NVIDIA V100 GPUs, resulting in almost negligible overhead. The new DPRc model and training procedure provide a potentially powerful new tool for the creation of next-generation QM/MM potentials for a wide spectrum of free-energy applications ranging from drug discovery to enzyme design.

1. INTRODUCTION

Classical molecular dynamics simulations using traditional molecular mechanics (MM) force fields have had a tremendous impact on life science applications.¹⁻³ However, the translation of this capability to the latest and most powerful state-of-the-art and emerging quantum mechanical (QM) methods has not kept pace.⁴ These high-level QM methods are extremely important for the study of catalytic mechanisms to guide enzyme design,⁵ drug discovery, and precision medicine applications.⁶ However, in practice, the computational cost of high-level ab initio quantum mechanical/ molecular mechanical (QM/MM)⁷⁻¹⁰ or quantum mechanical force field (QMFF)¹¹ simulations prohibits their practical use for many important applications. An attractive alternative is to use approximate QM methods such as the semiempirical¹² dorbital modified neglect of diatomic overlap $(MNDO/d)^{13}$ or density-functional tight binding (DFTB)^{14,15} methods that are typically 100-1000 times faster than ab initio QM. For large systems, these methods can be made both much faster (in terms of scaling) and more accurate for condensed-phase simulations using a linear-scaling quantum mechanical force field (QMFF) framework.^{11,16} Although some promising progress has been recently made,¹⁷ a critical barrier to progress for both QM/MM and QMFF methods has been in the design of robust models for the interactions between QM and MM regions or between QM fragments that provide the quantitative accuracy demanded by biocatalysis, drug discovery, and precision medicine applications.

Machine learning (ML)-based potentials afford a promising solution to the development of next-generation molecular simulation force fields with the efficiency comparable to that of MM force fields and accuracy that has started to approach that of high-level QM methods.¹⁸ In the past decade, researchers

Received: February 26, 2021 Published: October 13, 2021





have developed a broad spectrum of different ML potentials.¹⁹⁻³⁵ Recently, an ML-based model called deep potential-smooth edition (DeepPot-SE)³⁶ was developed to efficiently represent organic molecules, metals, semiconductors, and insulators with an accuracy comparable to that of *ab* initio QM models. The DeepPot-SE model has recently been highlighted in simulations of interfacial processes in aqueous aerosols³⁷ and large-scale combustion reactions in the gas phase³⁸ and demonstrated great success in providing predictive insight into complex reaction processes. To improve the accuracy and transferability of the DeepPot-SE models, the deep potential GENerator (DP-GEN) scheme^{39,40} uses an active-learning algorithm to generate models in a way that minimizes human intervention and reduces the computational cost for data generation and model training. The DP-GEN scheme has been successful in modeling metallic systems,^{39,40} chemical reactions at the interface of water and TiO_{2} ,⁴¹ transition from molecular to ionic ice at high pressure,⁴² gasphase reactive systems,⁴³ etc. These methods have evolved into an open-source software platform (DeePMD-kit⁴⁴ and DP-GEN⁴⁰) and have been enhanced with GPU acceleration and applied to simulations of 100 million atoms.^{45,46} Despite the achievement of these preliminary applications, in practice, pure ML potentials alone have had only limited success for condensed-phase MD simulations. $^{47-52}$ This is due to the fact that ML potentials often explicitly model short-ranged interactions without an explicit treatment of long-ranged interactions that are critical for modeling heterogeneous systems in the condensed phase.⁵³ In the present work, we develop a QM/MM model strategy that departs from a fast approximate QM method and uses ML potential as a correction term to greatly enhance the QM/QM and QM/ MM interactions so as to achieve much higher quantitative accuracy. The fast QM methods provide a robust model for short-ranged bonding and enable rigorous and efficient modeling of long-range many-body polarization and electrostatic interactions, which is the part of the interatomic interactions where pure ML potentials are unreliable. However, where the fast QM methods lack quantitative accuracy is in the short- or mid-range interactions where the ML correction potentials are exceptional. It is noteworthy that Kroonblawd, Goldman, and co-workers have recently reported promising success in generating accurate free-energy surfaces for chemical reactions using a similar-spirited approach using a semiempirical QM model and force matching^{17,54} to improve the interactions between QM atoms. Other recent works have applied force matching to collective variables within QM/MM simulations to calculate free-energy surfaces with an efficient semiempirical method.⁵⁵ In this approach, the interactions between QM and MM atoms are not explicitly corrected, but the reproduction of the net mean forces in the space of the reaction coordinate collective variables describing the reaction pathway implicitly accounts for their effect within the scope of the parameterization.

The ultimate goal is to develop new QM/MM and QMFF models based on high-level *ab initio* QM data in environments that mimic the condensed phase. In particular, we are interested in models that are able to reproduce free-energy (potential of mean force) surfaces and reaction pathways. The goal of the current work is to develop a robust training procedure for the ML neural networks such that the resulting models can robustly reproduce free-energy profiles from a reference model. Recall, however, that for high-level *ab initio*

pubs.acs.org/JCTC

QM methods, we cannot generally afford to simulate the QM/ MM reference free-energy profiles, at least not to sufficiently high precision to allow assessment of the new models. After all, this is the reason we wish to develop ML correction potentials that can be used along with an affordable QM/MM model for practical applications. Hence, in the present work, instead of using a high-level *ab initio* free-energy profile, which would not be feasible to compute, we examine two different approximate semiempirical QM models, MNDO/d¹³ and 2nd-order DFTB,¹⁴ that are substantially different in the functional form and also in terms of their predicted free-energy profile results to alternatively serve as baseline and target models. Specifically, we examine QM/MM simulations of a series of nonenzymatic model phosphoryl transfer reactions in aqueous solution (Figure 1).



Figure 1. Model nonenzymatic phosphoryl transfer reaction in solution. In a native RNA system, the positions labeled "X" correspond to oxygen positions (the canonical RNA numbering scheme is used to identify atomic positions). However, chemically modified variants involving thio substitution at one or more of these positions are commonly used in experimental mechanistic studies of RNA-cleaving enzymes.⁶⁵ In addition to the native model system, several of these variants will be studied in the present work (Table 1). The phosphoryl transfer reaction coordinate, $\xi_{\rm PT} = R_{\rm XS'-P} - R_{\rm X2'-P}$, is a difference between bond-breaking and bond-forming bond lengths, $R_{\rm XS'-P}$ and $R_{\rm X2'-P}$, respectively.

As both QM models are sufficiently affordable, we can determine high-level reference free-energy profiles from which to evaluate the performance of the ML correction potentials and validate the training procedures. The validated training procedures can then be applied in future work to develop ML correction potentials using high-level *ab initio* QM reference data, where it is not affordable to rigorously compute the high-level free-energy surfaces. As will be discussed below, different QM models affect not only the internal QM energy and forces but also those arising from QM/MM interactions. This brought to light limitations in the current DeepPot-SE model for QM/MM simulations and the need to develop an affordable solution to improving the QM/MM interactions as well as the internal QM energy and forces.

To achieve this, we have extended the DeepPot-SE model to include a range-corrected QM/MM interaction term that is smooth across the range boundary. This deep potential—range correction (DPRc) model is shown to be critical for correctly reproducing QM/MM reference data where the QM base and target models are significantly different. We further enhance the latest AMBER20⁵⁶ with our QM/MM and QMFF framework¹¹ interfaced with DeePMD-kit⁴⁴ and DP-GEN⁴⁰ to enable ML corrections for high accuracy. To demonstrate the accuracy and transferability of the approach using these efficient tools, we studied the 2'-O-transphosphorylation RNA cleavage reactions⁵⁷ (Figure 1). This specific phosphoryl

transfer (PT) reaction is essential for living organisms and has been the focus of many structural and mechanistic computational enzymology studies of nucleic acid enzymes known to catalyze it,^{5,58-61} yielding insight into nucleic acid enzyme design.⁶² This approach builds upon our recent work to develop new methods for construction of multidimensional free-energy surfaces⁶³ and "quantum mechanical book-ending" methods⁶⁴ for alchemical free-energy simulations, adding to the arsenal of computational tools for free-energy prediction in the AMBER suite of programs.

The paper is outlined as follows: the Methods section describes details of the DPRc models, active learning workflow, and computational details of the free-energy simulations. The Results and Discussion section first outlines the research strategy and specific objectives of the paper and then explores the effects of the QM/MM interaction range cutoff and initial data set on a native nonenzymatic phosphoryl transfer model system in solution. Next, we examine the effect of the active learning workflow on the transferability of the ML-corrected models outside the scope of their training by considering reactions of five additional chemically modified variants that represent thio substitutions used experimentally in mechanistic studies of RNA-cleaving enzymes.⁶⁵ Finally, we examine the behavior of the ML QM/MM interaction potential as well as the computational performance on both CPUs and GPUs. The paper concludes with a summary of key results and an outlook on future work.

2. METHODS

This section describes details of the neural networks and deep potential (DP) model and our extension to include smooth range corrections for QM/MM interactions. These methods have been implemented into the DeePMD-kit software,⁴⁴ and neural network training was facilitated by the DP-GEN software.⁴⁰ Combined QM/MM and QM/MM/DPRc simulations were performed with the sander program within AMBER20,⁵⁶ with interfaces built with DeePMD-kit to integrate the DPRc term into the potential energy. Analysis of free-energy surfaces was made using the variational free-energy profile (vFEP) method^{66,67} that has recently been extended to enable analysis of high-dimensional free-energy profiles⁶³ and implemented into the FE-ToolKit software package.⁶⁸ A more detailed description of each of these methods, along with computational details, is provided below.

2.1. Neural Networks. A neural network consists of an input layer, an output layer, and a sequence of N_{hidden} hidden layers connecting the input to the output. In the context of the present work, each layer is an array, and the length of the array is the number of nodes (neurons) within the layer. The M_k nodes of layer k are denoted as $\mathcal{L}_i^{(k)}(x)$ ($i = 1, \dots M_k$), where the input array, x, is all M_{k-1} node values from the previous layer or all M_0 node values of the input layer. The node values are computed from a biased, weighted sum of the inputs, where the weight W and bias b values are the neural network parameters to be trained. The dimensionality of the parameter arrays depends on the dimensionality of the inputs and outputs.

For example, the M_k nodes of layer k can be expanded from a scalar quantity x

$$\mathcal{L}_i^{e,(k)}(x) = \tanh(xW_i^{e,(k)} + b_i^{e,(k)}) \text{ for } \qquad 1 \le i \le M_k$$
(1)

pubs.acs.org/JCTC

The superscript "e" refers to the "expanded" size of the output. An array input can be expanded similarly

$$\mathcal{L}_{i}^{e,(k)}(\boldsymbol{x}) = \tanh\left(\sum_{j}^{M_{x}} x_{j}W_{ji}^{e,(k)} + b_{i}^{e,(k)}\right) \text{for} \qquad 1 \le i \le M_{k}$$
(2)

In this case, the weight matrix is a $M_x \times M_k$ matrix rather than an array of length M_k , where M_x is the length of the input array x.

A scalar output layer can be contracted from an input array by performing a weighted sum of all node values. We will denote this with a superscript "c" for scalar "contraction"

$$\mathcal{L}^{c,(k)}(\mathbf{x}) = \sum_{j}^{M_{x}} x_{j} W_{j}^{c,(k)} + b^{c,(k)}$$
(3)

Layer k could have twice as many nodes as the input layer, $M_k = 2M_x$, which shall be denoted with a superscript "d" for a "doubling" of size

$$\mathcal{L}_{i}^{d,(k)}(\mathbf{x}) = \begin{cases} x_{i} + \tanh\left(\sum_{j}^{M_{x}} x_{j}W_{ji}^{d,(k)} & \text{if } 1 \leq i \leq M_{x} \\ + b_{i}^{d,(k)}\right) \\ x_{i-M_{x}} & \text{if } M_{x} + 1 \leq i \leq 2M_{x} \\ + \tanh\left(\sum_{j}^{M_{x}} x_{j}W_{ji}^{d,(k)} \\ + b_{i}^{d,(k)}\right) \end{cases}$$

$$(4)$$

Finally, consecutive layers may have the same number of nodes, $M_k = M_{xx}$ in which case we use a superscript "*u*" to denote an "unchanged" size

$$\mathcal{L}_{i}^{u,(k)}(\boldsymbol{x}) = x_{i} + \tanh\left(\sum_{j}^{M_{k}} x_{j}W_{ji}^{u,(k)} + b_{i}^{u,(k)}\right) \text{for}$$

$$1 \le i \le M_{k}$$
(5)

2.2. Deep Potential (DP). The DP is a sum of atomic contributions to the energy

$$E = \sum_{i=1}^{N} E_i \tag{6}$$

The expression for the atomic contribution E_i is a neural network consisting of three hidden layers (see eq 7). The input layer is the "descriptor" array $D(\tilde{\mathcal{R}}_i)$ (see eq 8), which provides a description of the environment, inferred from the "environment matrix" $\tilde{\mathcal{R}}_i$ (see eq 9), the "embedding" matrix \mathcal{G}_i (see eq 11), and a reduced dimension embedding matrix $\mathcal{G}_i^<$ (see eq 12)

$$E_{i} = \mathcal{L}^{c}(\mathcal{L}^{u,(E_{i},3)}(\mathcal{L}^{u,(E_{i},2)}(\mathcal{L}^{e,(E_{i},1)}(D(\widetilde{\mathcal{R}}_{i})))))$$
(7)

$$D(\hat{\mathcal{R}}_i) = (\boldsymbol{\mathcal{G}}_i^{<})^T \cdot \hat{\mathcal{R}}_i \cdot (\hat{\mathcal{R}}_i)^T \cdot \boldsymbol{\mathcal{G}}_i$$
(8)

$$(\widetilde{\mathcal{R}}_{i})_{ja} = \begin{cases} s(R_{ij}), & \text{if } a = 1 \\ s(R_{ij})X_{ij}/R_{ij}, & \text{if } a = 2 \\ s(R_{ij})Y_{ij}/R_{ij}, & \text{if } a = 3 \\ s(R_{ij})Z_{ij}/R_{ij}, & \text{if } a = 4 \end{cases}$$
(9)

$$(\mathcal{G}_{i})_{jk} = \mathcal{L}_{k}^{d,(\mathcal{G}_{p}3)}(\mathcal{L}^{d,(\mathcal{G}_{p}2)}(\mathcal{L}^{e,(\mathcal{G}_{p}1)}(s(R_{ij}))))$$
(11)

$$(\mathcal{G}_i^{<})_{jk} = (\mathcal{G}_i)_{jk} \text{ for } 1 \le j \le N_{\text{neigh}}, \ 1 \le k \le M_{<}$$
(12)

Equation 10 is a switched reciprocal distance function that controls the range of the environment to be described. If an atom is separated from atom *i* by a distance greater than R_{off} , then the atom is not included in the environmental description of atom *i*. If a neighboring atom is within a distance of R_{on} , then the neighbor is given full weight in the description. The weight smoothly changes between these limits. Consequently, the environment matrix is a $N_{\text{neigh}} \times 4$ array, where N_{neigh} is the number of atoms within R_{off} .

Equation 11 is a second neural network; therefore, we use the superscripts (E_i, k) and (G_i, k) to refer to the layers appearing in eqs 7 and 11, respectively. The number of nodes in each layer is a choice that completes the definition of the neural network. The chosen number of nodes also defines the array sizes for several of the arrays used in the above expressions. In the present work, the three layers appearing in eq 7 use $M_1^{(E_i)} = 240$, $M_2^{(E_i)} = 240$, and $M_3^{(E_i)} = 240$ nodes. The three layers appearing in eq 11 use $M_1^{(G_i)} = 25$, $M_2^{(G_i)} = 50$, and $M_3^{(G_i)} = 100$ nodes. Each row of the embedding matrix corresponds to a neighbor, so the size of the matrix is $N_{\text{neigh}} \times M_3^{(G_i)}$.

The reduced dimensional embedding matrix has the same values as \mathcal{G}_i ; however, only the first $M_<$ columns are stored, where $1 \leq M_< \leq M_3^{(\mathcal{G}_i)}$. In other words, the dimensions of $\mathcal{G}_i^<$ are $N_{\text{neigh}} \times M_<$. Small values of $M_<$ will increase computational

pubs.acs.org/JCTC

performance but decrease numerical accuracy. One empirically chooses the smallest value of $M_{<}$ that yields acceptable accuracy. We have found $M_{<}$ = 12 to be an acceptable value in the present work.

The values of the descriptor array can be viewed as elements of a $M_{<} \times M_{3}^{(\mathcal{G}_{i})}$ matrix (eq 11); however, it is viewed as a vector of length $M_{<} \cdot M_{3}^{(\mathcal{G}_{i})}$ when used as the input layer in eq 7.

2.3. Range-Corrected Deep Potential (DPRc). We seek to develop a computationally efficient approximation to replace costly ab initio Hamiltonian evaluations within OM/MM MD simulations. There are several ways that one might approach this. Some researchers might prefer to eliminate the use of physics-based models entirely and solely relying on an MLtrained model; however, current pure ML models do not explicitly contain a treatment for long-range electrostatic interactions, which are important to maintain stable RNA and DNA secondary structures, for example. The approach explored in this work is to substitute the ab initio QM Hamiltonian with a semiempirical Hamiltonian and use an ML model to correct short-range atomic interactions to mimic those obtained with an expensive Hamiltonian. We are motivated to use a semiempirical Hamiltonian as a foundation for correction because the physics built in to semiempirical Hamiltonians explicitly models long-range interactions, and they offer a reasonable description of bond making and breaking events at an affordable cost.

The QM/MM calculations considered in this work use electrostatic embedding; therefore, the QM electrostatic interaction with the MM region changes when replacing the ab initio Hamiltonian with a semiempirical model. The largest discrepancies between the electrostatic interactions will be between the QM region and the nearby MM residues. In other words, to make the semiempirical QM/MM method mimic the ab initio QM/MM energies and forces, the ML correction potential needs to modify the interactions between the semiempirical atoms with the other semiempirical atoms and the nearby MM atoms. Alternatively, replacing the QM Hamiltonian has no effect on the MM interactions with other MM atoms. The DP can be applied to the improvement of QM/QM interactions; however, a few changes need to be made for it to include correct QM/MM interactions without affecting MM/MM interactions in a manner that conserves energy as MM residues diffuse into (or out of) the vicinity of the QM region. The modified ML model shall be referred to as the range-corrected deep potential, which we abbreviate as DPRc and describe below.

The electrostatically embedded QM/MM energy with a DPRc ML potential is shown in eq 13

$$E(\mathbf{R}; \mathbf{P}) = E_{\text{QM}}(\mathbf{R}; \mathbf{P}) + E_{\text{QM}/\text{MM}}(\mathbf{R}; \mathbf{P}) + E_{\text{MM}}(\mathbf{R}) + E_{\text{ML}}(\mathbf{R})$$
(13)

R is an $N \times 3$ array of atomic coordinates, and **P** is the QM Hamiltonian's single-particle density matrix. $E_{\rm QM}(\mathbf{R}; \mathbf{P})$ and $E_{\rm MM}(\mathbf{R})$ are the QM and MM energies, respectively. $E_{\rm QM/MM}(\mathbf{R}; \mathbf{P})$ contains the electrostatic and Lennard-Jones (or other nonelectrostatic-nonbonded model) interactions between the QM and MM regions. $E_{\rm ML}(\mathbf{R})$ is the DPRc ML potential.

If the ML correction was applied only to the QM atoms, then one could train the existing deep potential model to obtain parameters for each atomic number, and the resulting

model would correct the energies and the forces between the QM atoms. When considering a layer of MM atoms, some modifications are necessary. The correction applied to the interaction between an MM hydrogen and the QM atoms will be much different from the correction between a QM hydrogen and the other QM atoms. In other words, ML weight and bias parameters are trained for QM atomic numbers and MM force field "atom types". One can imagine other schemes, such as training the ML parameters for MM atoms using the MM atomic charge as an additional input; however, the MM atomic charge information is partially encoded within the force field atom type assignments.

The interaction between MM atoms with other MM atoms remains unchanged upon replacing the *ab initio* Hamiltonian with a semiempirical model. The ML correction should not alter the MM/MM interactions. During the course of an MD simulation, the MM residues may diffuse away from (or toward) the QM region. To maintain energy conservation, the ML correction to the QM/MM interactions must approach 0 as the MM atoms leave the area of the QM region, as defined by a cutoff distance. Furthermore, the MM atoms cannot contribute a (one-body) constant to the ML correction to the energy.

To satisfy these requirements, the definition of the switched reciprocal function (eq 10) is modified to eliminate MM atoms from the environment of other MM atoms and formally will have an explicit dependence on the atom indexes in addition to their separation distance

$$s_{ij}(R_{ij}) = \begin{cases} 0, & \text{if } ij \in \text{MM} \\ \frac{1}{R_{ij}}, & \text{if } ij \in \text{QM} \\ \text{if } i \in \text{QM} \land j \in \text{MM} \text{ or } i \in \text{MM} \land j \in \text{QM}: \\ \frac{1}{R_{ij}}, & \text{if } R_{ij} \leq R_{\text{on,QM/MM}} \\ \frac{1}{r} \left\{ \left(\frac{r - R_{\text{on}}}{R_{\text{off}} - R_{\text{on}}} \right)^3 & \text{if } R_{\text{on,QM/MM}} \\ -6 \left(\frac{r - R_{\text{on}}}{R_{\text{off}} - R_{\text{on}}} \right)^2 & < R_{ij} \\ \left(-6 \left(\frac{r - R_{\text{on}}}{R_{\text{off}} - R_{\text{on}}} \right)^2 & < R_{\text{off,QM/MM}} \\ + 15 \frac{r - R_{\text{on}}}{R_{\text{off}} - R_{\text{on}}} \\ - 10 \right) + 1 \right\}, \\ 0, & \text{if } R_{ij} \geq R_{\text{off,QM/MM}} \end{cases}$$

$$(14)$$

This definition creates different R_{on} and R_{off} parameters for QM/QM interactions and QM/MM interactions, so we may explore the inclusion of QM/MM corrections without disturbing the QM/QM interactions. Furthermore, the definition of the atomic contribution must be modified to avoid one-body contributions from the MM atoms

pubs.acs.org/JCTC

$$E_{i} = \begin{cases} \mathcal{L}^{\epsilon}(\mathcal{L}^{u,(E_{i},3)}(\mathcal{L}^{u,(E_{i},2)}(\mathcal{L}^{e,(E_{i},1)}(\mathbf{D}(\mathcal{R}_{i}))))), & i \in \text{QM} \\ \mathcal{L}^{\epsilon}(\mathcal{L}^{u,(E_{i},3)}(\mathcal{L}^{u,(E_{i},2)}(\mathcal{L}^{e,(E_{i},1)}(\mathbf{D}(\mathcal{R}_{i}))))) & i \in \text{MM} \\ - E_{i}^{(0)}, \end{cases}$$
(15)

 $E_i^{(0)}$ is the one-body contribution to the energy

$$E_{i}^{(0)} = \mathcal{L}^{c}(\mathcal{L}^{u,(E_{i},3)}(\mathcal{L}^{u,(E_{i},2)}(\mathcal{L}^{e,(E_{i},1)}(0))))$$
(16)

2.4. Computation Details. We seek to use an ML-corrected semiempirical model as a replacement for an *ab initio* Hamiltonian within QM/MM simulations; however, a robust training procedure for the ML correction must be developed. The development requires answering several questions, such as

- To what extent does "active learning" improve the fitted model ability to reproduce free-energy profiles?
- Do the models produce better free-energy profiles if the training data include samples from high-temperature ensembles?
- What is an acceptable cutoff for correcting the interaction between QM and MM regions?
- Do the models reproduce free-energy profiles for systems not included in the training?

Answering these questions requires extensive *ab initio* QM/ MM simulation sampling. Our approach to developing a training procedure is to use the ML model to mimic the energies and forces of another semiempirical Hamiltonian. For example, we can extensively sample free-energy surfaces with MNDO/d and then create an ML-corrected DFTB2 that reproduces the MNDO/d forces and compare the resulting free-energy profiles. In addition, we could create a reference DFTB2 profile, train an ML-corrected MNDO/d model, and compare the profiles.

For the purpose of developing the training procedure, we create free-energy profiles of the model nonenzymatic phosphoryl transfer reaction (Figure 1) in an explicit solvent. The native state contains oxygens at the 2', 3', 5', and nonbridging phosphoryl positions. We shall also explore various thio-substituted model compounds (see Table 1) to

Table 1. Reactions of Native System and Chemically Modified Variants^a

reaction	X2′	X5′	XP2	XP1	X3′
native	0	0	0	0	0
O2'/S (S2')	S	0	0	0	0
O5'/S (S5')	0	S	0	0	0
O3'/S (S3')	0	0	0	0	S
OP1/S (S1P)	0	0	0	S	0
OP2/S, OP1/S (S12)	0	0	S	S	0

"Reactions of different modified variants were simulated, where the phosphorus atom has up to five bonds to different atoms during the course of the reaction (for chemical illustration, see Figure 1). X2' and X5' refer to the nucleophile and leaving group positions, respectively; XP2 and XP1 refer to the pro- R_P and pro- S_P nonbridge phosphoryl positions, respectively; S12 refers to a double thio substitution at the nonbridge positions; X3' refers to the bridging position to the C3' of the pseudo-sugar ring. Here "X" refers generically to either oxygen "O" (as in the native reaction) or sulfur "S" (for a chemically modified non-native reaction). The abbreviations for the reaction used elsewhere in the figures, tables, and text are indicated in parentheses.

pubs.acs.org/JCTC

Article

Table 2.	Summary of	f Different N	IL Parameter	Sets That	Vary Bas	e and '	Target	QM Models,	QM and	QM/MM	Interaction
Cutoffs,	and Initial I	Data Sets ^a									

ML	QM base + ML \rightarrow QM target	$R_{\rm C}/R_{\rm QM}$ (Å)	initial data set	data size
1a	$MNDO/d + ML \rightarrow DFTB2$	$0.0/\infty$	TREMD@298,315,330	6284 (1.17%)
2a	$MNDO/d + ML \rightarrow DFTB2$	3.0/∞	TREMD@298,315,330	6305 (1.18%)
3a	$MNDO/d + ML \rightarrow DFTB2$	$6.0/\infty$	TREMD@298,315,330	6369 (1.19%)
4a	$MNDO/d + ML \rightarrow DFTB2$	9.0/∞	TREMD@298,315,330	6370 (1.19%)
5a	$MNDO/d + ML \rightarrow DFTB2$	6.0/6.0	TREMD@298,315,330	5970 (1.11%)
6a	$MNDO/d + ML \rightarrow DFTB2$	$6.0/\infty$	TREMD@298,(315),(330)	3354 (0.75%)
7a	$MNDO/d + ML \rightarrow DFTB2$	$6.0/\infty$	MD@298	2615 (0.59%)
8a	$MNDO/d + ML \rightarrow DFTB2$	$6.0/\infty$	TREMD/TGT@298,315,330	4429 (0.83%)
1b	DFTB2 + ML \rightarrow MNDO/d	$0.0/\infty$	TREMD@298,315,330	6170 (1.15%)
2b	DFTB2 + ML \rightarrow MNDO/d	3.0/∞	TREMD@298,315,330	4617 (0.86%)
3b	DFTB2 + ML \rightarrow MNDO/d	$6.0/\infty$	TREMD@298,315,330	4697 (0.88%)
4b	DFTB2 + ML \rightarrow MNDO/d	9.0/∞	TREMD@298,315,330	4415 (0.82%)
5b	DFTB2 + ML \rightarrow MNDO/d	6.0/6.0	TREMD@298,315,330	4232 (0.79%)
6b	DFTB2 + ML \rightarrow MNDO/d	$6.0/\infty$	TREMD@298,(315),(330)	2299 (0.51%)
7b	DFTB2 + ML \rightarrow MNDO/d	$6.0/\infty$	MD@298	2267 (0.51%)
8b	DFTB2 + ML \rightarrow MNDO/d	$6.0/\infty$	TREMD/TGT@298,315,330	4617 (0.86%)

^{*a*}The notation (QM base) + ML \rightarrow (QM target) is used to represent the development of an ML correction to the QM base model in order to reproduce reference results from the target QM model. Models consider different cutoffs, $R_C \equiv R_{off,QM/MM}$, for the smooth range correction for QM/MM interactions (a value of 0 indicates no range correction), and $R_{QM} \equiv R_{off,QM}$ for the internal QM interactions (a value of ∞ indicates no cut-off). The value for $R_{on,QM/MM}$ in the switched reciprocal function in eq 14 was chosen to be fixed to 1 Å (except for $R_C = 0$). The value of $R_{on,QM} = \infty$). These values were chosen to be consistent with previous works using the DP model.^{38,43} In addition, we consider different initial data used in the ML potential training. Specifically, we consider (1) "MD@298", traditional MD at 298 K, (2) "TREMD@298,(315),(330)", enhanced sampling with TREMD at 298, 315, and 330 K but data only collected at 298 K, (3) "TREMD@298,315,330", TREMD using data at 298 K in addition to data from elevated temperatures of 315 and 330 K, and (4) "TREMD/TGT@ 298,315,330", the same as "TREMD@298,315,330" but performing simulations using the QM target model. All training (that required re-evaluation at the target QM level) is shown in the rightmost column, as well as the percentage of selected data points relative to the total data sets marked in bold (3a and 3b) were ultimately selected to be applied to develop and test robust DPRc models trained for a broader set of variant reactions.

explore the transferability of the trained models. That is, MLcorrected Hamiltonians will be trained using the energies and forces gathered from a subset of the thio-substituted systems, and the trained models will be used to calculate the free-energy profiles of the systems that did not contribute to the set of training data.

2.4.1. Simulation Setup. The initial structure for each RNAlike nonenzymatic model system shown in Figure 1 was generated from SMILES with Openbabel⁶⁹ and then parameterized using the AMBER ff14SB/GAFF force field70,71 together with partial charges generated by AM1-BCC.⁷² The entire RNA-like solute molecule was defined as the QM region (QM solute). For each system, the QM solute was solvated with 1510 TIP4P/Ew waters.73 Simulations were performed with the sander program within AMBER20⁵⁶ using a 1 fs time step. The equilibration procedure consisted of 100 ps of heating using a Langevin thermostat to increase the temperature from 0 to 298 K.⁷⁴ The system density was equilibrated in the isothermal-isobaric ensemble at 1 atm and 298 K. Pressure was regulated with a Berendsen barostat using a 5 ps collision frequency.⁷⁵ The Lennard-Jones potential was truncated at 8 Å, and a long-range tail correction is used to model the interactions beyond the cutoff. The long-range electrostatics was evaluated with the particle mesh Ewald method using a 1 Å³ grid spacing.^{9,76}

The RNA cleavage reaction⁵⁷ is described by the reaction coordinate (progress variable) ξ_{PT} defined as

$$\xi_{\rm PT} = R_{\rm X5'-P} - R_{\rm X2'-P} \tag{17}$$

where $R_{X5'-P}$ and $R_{X2'-P}$ are the bond-breaking and bondforming bond lengths in Figure 1, respectively. After equilibration of the reactant structure, umbrella window simulations were prepared to scan the reaction coordinate $\xi_{\rm PT}$ from -4 to 5 Å in steps of 0.1 Å (91 umbrella windows in total). The initial coordinates of each window were generated by sequentially equilibrating each umbrella window for 25 ps in the canonical ensemble at 298 K starting from the nearest available equilibrated structure. Free-energy profiles were analyzed from 100 ps of canonical ensemble sampling at 298 K. The 100 ps of sampling for each window was generated from four independent 25 ps simulations initiated from different random number seeds. The vFEP method,^{66,67} which has recently been extended for analysis of highdimensional free-energy profiles⁶³ and implemented into the FE-ToolKit software package,⁶⁸ was used to analyze the simulation results to generate the free-energy profiles.

2.4.2. Neural Network Training. The neural network parameters are optimized to reproduce a target set of energies and forces. If the DFTB2 QM/MM system is the target, then the ML model is trained to reproduce the difference between the DFTB2 QM/MM and MNDO/d QM/MM energies and forces. Similarly, if the MNDO/d QM/MM system is the target, then the target quantities are the differences between the MNDO/d QM/MM and DFTB2 QM/MM energies and forces. The energies and forces include all atoms in the system. The training is improved by providing the optimization algorithm more samples, and the efficiency of the optimization can be improved by using active learning (AL).

2.4.3. Active Learning (AL) Workflow. AL is a procedure whereby the set of target data is increased to improve the fit; however, the new data are selected to avoid including redundant information within the existing training set. The three steps in an AL cycle are training, exploration, and labeling. The training phase parameterizes multiple ML models from the existing collection of target data. The ML training is a stochastic procedure, and the multiple ML models are obtained by initiating the fit procedure from different random seeds. In the present work, we generate four ML models per fit. The exploration process uses one of the new ML models to sample phase space. In the current context, this means that we

perform a new set of ML-corrected semiempirical QM/MM umbrella window simulations using one of the four ML models. The trajectories are then reanalyzed to calculate the ML-corrected QM/MM energies and forces for each of the four ML models; if the four sets of energies and forces sufficiently agree, then the trajectory frame is discarded. The selected frames are chosen from the trajectory according to the condition

$$\{\mathcal{R}_{n}|n \in I_{\text{cand}}, \qquad I_{\text{cand}} = \{n|\theta_{\text{low}} \le \varepsilon_{n} < \theta_{\text{high}}\}\}$$
(18)

where the model deviation ϵ , proposed in the previous work,⁴⁰ is the maximal standard deviation of the atomic force predicted by four models. In this work, θ_{low} and θ_{high} are set to 0.1 and 0.25 eV/Å, respectively, as the model deviation of most points in the training data is less than 0.1 eV/Å. To reduce the computational cost, if the number of selected frames is larger than 1000, then 1000 of these frames are chosen at random. The labeling phase evaluates the target energies and forces of the selected frames chosen within the exploration step. The newly analyzed data are included as additional reference data used to train the network parameters in subsequent active learning cycles. In this work, we performed nine AL cycles for each ML parameter set and confirmed that the ratio of accurate frames ($\epsilon < \theta_{low}$) in all final simulations (see Supporting Information) is greater than 99.6%. The AL procedure is a valuable tool to efficiently expand the information contained in the training because the target energies and forces may be expensive to compute.

2.4.4. Initial Data. The energies and forces used to perform the fit (or the first fit in an AL procedure) can be analyzed from the trajectory frames of the umbrella window simulations. This is denoted by "MD@298" in Table 2. One could instead perform temperature replica exchange simulations (TREMD); in Table 2, the rows marked "TREMD@298, 315, 330" generate initial data from TREMD simulations involving three temperatures: 298, 315, and 330 K. The rows marked "TREMD@298, (315), (330)" perform TREMD simulations, but the training data only contain the samples taken from the 298 K ensemble. "TREMD/TGT@298, 315, 330" performs the same as "TREMD@298, 315, 330" but uses the target potential. The initial training data consist of 500 frames per umbrella window simulation. Each umbrella window simulations was run for 25 ps (1 fs time step), and each frame was stored every 50 fs. The reaction coordinate is divided into 91 umbrella windows, so the "MD@298" and "TREMD@298, (315), (330)" protocols each produce 45.5k frames for the initial training. The "TREMD@298, 315, 330" and "TREMD/ TGT@298, 315, 330" protocol produces 136.5k frames because all three temperatures are included in the training. However, to reduce the size of the initial data, we only use 1000 frames from "MD@298" or "TREMD@298, (315),

pubs.acs.org/JCTC

(330)" or 3000 frames from "TREMD@298, 315, 330" or "TREMD/TGT@298, 315, 330".

3. RESULTS AND DISCUSSION

3.1. Research Strategy. Herein, we develop a robust training procedure for the ML neural networks that can be used as corrections to a base QM model in order to robustly reproduce free-energy profiles from a target QM model. Such a procedure will be extremely important for developing nextgeneration fast QM force fields that are accuracy comparable to very high-level QM methods for which simulation is prohibitive due to excessive computational requirements. Hence, in order to develop and validate such a robust procedure, we chose two fast approximate QM models, and for both, it is possible to compute benchmark reference freeenergy profiles for a series of chemical reactions: MNDO/d¹³ and second-order DFTB.¹⁴ These QM models make approximations that enable greatly enhanced computational efficiency relative to high-level ab initio QM methods and take recourse into empiricism to recover considerable accuracy. These models are also significantly different from one another in terms of the approximations they make and consequently their functional form. These models were intentionally chosen because they are known to give significantly different reaction profiles for the phosphoryl transfer reactions in the present study.^{77,78} In this way, the goal of developing a robust training procedure for an ML potential that "corrects" one QM model into the other presents a stringent stress test.

Both MNDO/d and DFTB2 allow for d-orbitals in their representation of third-row elements, which has been shown to be important for modeling phosphoryl transfer reactions⁷⁹⁻ as well as bonding of sulfur atoms in some cases. The MNDO/ d method is based on the neglect of diatomic overlap, which enables a framework for electrostatics to be modeled as a set of atom-centered multipoles. However, the overlap matrix used in the eigenvalue problem that must be solved self-consistently is assumed to be the unit matrix, and hence, the normal exchange repulsions that would normally arise from orthogonalization of the molecular orbitals are absent and must be modeled in a different way.⁸² The DFTB2 method uses a second-order density functional expansion and a two-center integral approximation that enables very fast computation of a firstorder effective Hamiltonian (tight-binding) matrix.⁸³ Electrostatic interactions are captured as a second-order term, which, like the MNDO/d method, demands a self-consistent procedure to solve. A rigorous atomic overlap matrix is used in the generalized eigenvalue problem that more naturally overcomes issues related to orthogonalization. However, this also introduces complications in the representation of electrostatic interactions, and in the DFTB2 model, these interactions are modeled as atomic monopoles with charges determined from density matrix partitioning.

As mentioned above, these QM methods were not chosen for their accuracy for the phosphoryl transfer reactions being used here but rather because their differences pose a stringent test for the purpose of training ML potentials as QM model corrections. In fact, phosphoryl transfer reactions are of such importance in biology that specialized models have been developed based on a semiempirical MNDO/d-like⁸⁴ and third-order DFTB^{85,86} that greatly improve the accuracy for these reactions. Of key importance to recognize is that the MNDO/d and DFTB2 methods have substantial differences in the way they model orbital orthogonalization and electro-

	errors (shown in parentheses)										
	MNDO/d +	$ML \rightarrow DFTB2$		DFTB2 + ML \rightarrow MNDO/d							
ML	$\xi^{\ddagger}_{ m PT}$	ΔG^{\ddagger}	ΔG	ML	$\xi^{\ddagger}_{ m PT}$	ΔG^{\ddagger}	ΔG				
1a	(-0.18)	(-2.8)	(-0.3)	1b	(0.05)	(-3.0)	(-9.3)				
2a	(-0.05)	(-1.7)	(0.0)	2b	(0.00)	(-0.6)	(0.5)				
3a	(-0.05)	(-2.4)	(0.2)	3b	(0.00)	(-0.8)	(0.1)				
4a	(0.00)	(-2.3)	(0.5)	4b	(0.00)	(-0.9)	(-0.5)				
5a	(0.05)	(-1.9)	(1.9)	5b	(0.00)	(-1.1)	(-0.6)				
6a	(0.00)	(-1.6)	(0.9)	6b	(0.00)	(-1.3)	(-1.0)				
7a	(0.00)	(-1.8)	(1.2)	7b	(0.00)	(-0.6)	(0.2)				
82	(-0.05)	(-26)	(1 1)	8h	(0.00)	(-10)	(0.1)				

Table 3. Comparison of Target and Model Transition-State Coordinate Values, ξ_{PT}^{\ddagger} (Å), Free-Energy Barriers, ΔG^{\ddagger} (kcal/mol), and Reaction Free Energies, ΔG (kcal/mol), for the Native Reaction with Different ML Parameter Sets^{*a*}

^aThe ML parameter sets are described in Table 2. These models were all trained on data only for the native reaction in order to identify the optimal training procedure to then apply to a broader set of variant reactions later. Reaction free-energy values were taken as the difference between the reactant state minimum and the value of the free-energy profile at 4 Å. Errors are shown in parentheses as the difference between target and model values. Details are provided in the text, and the full free-energy profile curves are shown in Figure 2. ML parameters sets 3a and 3b (marked in bold) were determined to be the best balance of accuracy and computational cost, and the corresponding QM and QM/MM interaction cutoffs and initial data set procedures (Table 2) were selected to be applied to develop and test robust DPRc models trained to a broader set of variant reactions.



Figure 2. Free-energy profiles of different parameters defined in Table 2, compared with original DFTB2 and MNDO/d curves. The left panels involve ML parameters that use MNDO/d as the base QM model and DFTB2 as the target QM reference, whereas the right panels use DFTB2 as the base QM model and MNDO/d as the target QM reference. The top panels compare range correction values of 3, 6, and 9 Å and QM range values of 6 Å and ∞ , all using initial data from TREMD at 298, 315, and 330 K. The bottom panels compare initial data from conventional MD and TREMD at 298 K and from TREMD using 298, 315, and 330 K as well as from data collected from the QM target, all using a range correction of 6 Å. All ML curves were calculated from four independent MD simulations at the last cycle of AL. All reference curves were calculated from four independent MD simulations at the last cycle of AL. All reference curves were calculated from four independent MD simulations at the last cycle of AL. All reference curves were calculated from four independent MD simulations at the last cycle of AL.

statics. The former implies that one can expect significant differences in short-ranged exchange repulsions, including 1–4 interactions of the QM atoms that effect rotation about single bonds and in particular the puckering of five-membered rings.^{87–89} The latter implies that there will also be differences in electrostatic interactions from different multipolar representations that affect interactions between QM and MM atoms in the mid-range.

In this way, we introduce a considerable challenge developing a robust ML training procedure able to correct one QM model into the other and vice versa. Development and validation of such a procedure would enable the design of new fast QM + ML models, built upon the latest, most advanced QM base models with ML correction potentials trained with high-level *ab initio* QM target reference data. It should be noted that the ML training procedure used here is designed to match the forces of the target QM model. Here,

we compare free-energy results derived from performing new simulations using the QM + ML models and constructing freeenergy profiles^{63,66,67} for the chemical reactions along 1D and later in the paper 2D coordinates. Hence, the free-energy profiles being compared were not themselves explicitly considered in the training, although reference force data were obtained from simulations along the reaction path. Agreement of the analyzed free-energy profile results is thus a sensitive indicator of the robustness of the training procedure that considered only force data.

In the sections that follow, we will use the notation (QM base) + ML \rightarrow (QM target) to indicate how the ML correction is being applied, and we will alternatively use MNDO/d and DFTB2 as QM base and target models. In the second subsection that follows, we will explore the effects of the QM/MM interaction range cutoff and initial data set on a native nonenzymatic phosphoryl transfer model system. In the third subsection, we develop a robust active learning workflow and test the robustness of the ML-corrected models on a series of reactions as well as their transferability outside the scope of their training. In the fourth subsection, we will illustrate the nature of the DPRc model in correcting short- and mid-range QM/MM interaction potential and examine more broadly the computational scaling on both CPUs and GPUs for different QM/MM interaction range correction cutoffs.

3.2. Effect of Range Correction and Initial Training Data. We examined the effect of varying the cutoff for the range correction, along with the initial data used in training, on the accuracy of the resulting parameter sets for the native reaction illustrated in Figure 1 (where the "native" reaction has oxygen for all the "X" atoms in the figure). Several parameter sets were developed and tested and are summarized in Table 2. We tested the degree to which the force accuracy on an atom might vary with the number of MM neighbors and found this variation to be negligible compared to the distribution of force errors (see Supporting Information).

Table 3 and Figure 2 compare the free-energy profiles for each of the trained parameter sets listed in Table 2. In each panel of the figure, the uncorrected MNDO/d (red) and DFTB2 (black) are shown in bold line. It is clear that the profiles are considerably different in shape, overall reaction free energy, and curvature around the reactant state minimum. The top panels (Figure 2a,b) compare QM/MM interaction range correction cutoff values of $R_{\text{off},\text{QM/MM}} = 3$, 6, and 9 Å and $R_{\rm off,OM} = 6$ and ∞ , all using initial data from TREMD at 298, 315, and 330 K (TREMD@298, 315, 330 protocol). As R_{off,OM/MM} increases, a larger radius of MM atoms is included in the DPRc correction potential. The bottom panel (Figure 2c,d), on the other hand, compares the effect of the initial training data set (prior to the active learning cycles) using a cutoff value of $R_{\text{off},\text{QM/MM}} = 6$ Å. The DPRc models were initially trained on data derived from one of the four protocols [MD@298, TREMD@298, (315), (330), TREMD@298, 315, 330, or TREMD/TGT@298, 315, 330]. Each model then underwent nine cycles of active learning, each of which produced several additional frames per umbrella window simulation that were then included in the labeling phases. The profiles within the figure were generated from final models after the active learning cycles.

It is clear that in the absence of a range correction $(R_{\text{off},\text{QM/MM}} = 0 \text{ Å})$, the DPRc potential does not lead to close agreement with the target QM results (particularly for the DFTB2 + ML model). Consideration of a range correction out

to 3 Å considerably improves the agreement. Extending the range correction to 6 Å leads to profiles that are visually almost indistinguishable from the target reference results (and is within the statistical uncertainty of the calculations; data not shown in order to maintain better clarity of the plot lines). Further extension out to 9 Å does not provide any significant additional benefit. It is also important to apply the interaction range of ∞ within the QM region.

While the DPRc models were sensitive to the range correction cutoff, they are less affected by the initial data set used prior to active learning optimization cycles (bottom panels of Figure 2). The only notable exception occurs for the MNDO/d + ML model using replica exchange data only at 298 K, and the overall reaction free energy (orange line, bottom left panel in Figure 2) is elevated and remains closer to the uncorrected MNDO/d result. Inclusion of data from higher temperatures alleviates this issue and provides overall the closest agreement with the target QM reference data. We recommend the use of enhanced sampling with TREMD for generation of initial data in order to broaden the conformational space considered. As this method also requires coupled simulations at higher temperatures, we further recommend that these data be used in the ML training. The philosophy here is that it is important to include representative data for not only low-energy conformations that are likely to be frequently sampled but also higher-energy conformations that should rarely be sampled at ambient temperatures. In this way, the ML correction potentials can be trained to appropriately occupy/ avoid low-energy/high-energy regions of conformational space, as appropriate.

Taken together, these are important results. The need to include a range correction out to 6 Å suggests that for QM/ MM simulations, it may not be enough to simply correct the internal QM energies and forces but that explicit consideration of ML corrections for the mid-range QM/MM interactions is also crucial. Further, it is prudent to include not only simulation data at ambient temperatures but also data from higher-temperature enhanced sampling simulations. Moving forward, we adopt the use of a 6 Å QM/MM interaction range correction and include initial data from TREMD simulations at 298, 315, and 330 K.

While our initial training data, in some cases, considered data from higher-temperature data from TREMD simulations, our AL procedure only considered new data from simulations at 298 K. It might be the case that fine tuning of the AL procedure by integration of higher-temperature TREMD data could lead to even more robust models. This is a topic that we intend to explore in the future work.

3.3. Transferability and ML Model Validation. To check the transferability of the model, we examine five different variants from the native reaction where one or more oxygen positions are replaced with sulfur (Figure 1). These variants, along with their abbreviations, are listed in Table 1. These substitutions are often used experimentally as probes in mechanistic studies of RNA-cleaving enzymes⁶⁵ and particularly RNA enzymes. Sulfur is larger, softer, less electronegative, and more polarizable than oxygen. Thiols tend to have pK_a values that are 4–5 units below that of corresponding alcohols, indicating that alkyl sulfides are more stable in solution than alkyl oxides. Hence, for example, an S5' variant is an "enhanced leaving group" that can be used to probe to rescue the effects of knocking out a presumed catalytic general acid. An S2' variant, on the other hand, is expected to be more

Table 4. Comparison of Target (MNDO/d) and Model (DFTB2 + ML) Transition-State Coordinate Values, ξ_{PT}^{\ddagger} (Å), Free-Energy Barriers, ΔG^{\ddagger} (kcal/mol), and Reaction Free Energies, ΔG (kcal/mol), for Native and Variant Reaction Models^{*a*}

	DFTB2 + ML \rightarrow MNDO/d												
	target values												
		MNDO/d		1	ML (no AL)		Ν	ML (with AL)			no ML (DFTB2)		
reaction	$\xi^{\ddagger}_{ m PT}$	ΔG^{\ddagger}	ΔG	$\xi^{\ddagger}_{ m PT}$	ΔG^{\ddagger}	ΔG	$\xi^{\ddagger}_{ m PT}$	ΔG^{\ddagger}	ΔG	$\xi^{\ddagger}_{ m PT}$	ΔG^{\ddagger}	ΔG	
native	0.61	20.0	-5.2	(0.05)	(0.1)	(-0.4)	(0.00)	(-0.3)	(-0.5)	(-0.32)	(1.6)	(-2.9)	
S2′	0.16	30.5	4.3	(0.05)	(2.8)	(1.3)	(0.00)	(-0.5)	(1.4)	(-0.23)	(23.8)	(46.9)	
S5′	-0.02	10.0	-18.5	(0.05)	(1.5)	(-1.4)	(0.05)	(0.5)	(0.4)				
*S3′	0.66	20.1	-5.5	(-0.05)	(-1.4)	(-1.7)	(-0.09)	(0.1)	(1.0)	(-1.31)	(-4.7)	(-8.7)	
*S1P	0.75	23.2	-3.1	(0.00)	(1.5)	(0.4)	(-0.05)	(0.5)	(0.7)	(-2.04)	(-19.8)	(-15.3)	
S12	0.79	28.3	-1.2	(0.00)	(2.7)	(-1.0)	(0.05)	(-0.9)	(-1.4)	(0.59)	(-13.1)	(-4.7)	
ME				(0.02)	(1.2)	(-0.5)	(-0.01)	(-0.1)	(0.3)	(-0.84)	(-3.7)	(-4.8)	
MAE				(0.03)	(1.7)	(1.1)	(0.04)	(0.5)	(0.9)	(1.03)	(12.2)	(20.5)	

^{*a*}Native and variant model reactions are described in Table 1. The models were trained on data for the native reaction as well as S2', S5', and S12 variants. Note: reactions of variants indicated by an asterisk (*) were NOT used in the ML training and are examined to test the transferability of the ML correction to reactions outside the training scope. Reaction free-energy values were taken as the difference between the reactant state minimum and the value of the free-energy profile at 4 Å. The target QM reference is MNDO/d, and the model is DFTB2 + ML correction-trained using the procedures for ML parameter set 3b in Table 2. Errors are shown in parentheses as the difference between target and model values. Two ML correction potentials are compared. The first, designated as ML (no AL), is trained on an initial set of simulation data with no active learning. The second, designated as ML (with AL), uses the AL procedure to further evolve and improve the DL potential. In addition, errors with respect to the uncorrected (DFTB2) values are shown. The last two rows list mean error (ME) and mean absolute error (MAE). Details are provided in the text, and the full free-energy profile curves are shown in Figure 3.

reticent toward nucleophilic attack to phosphorus.⁹⁰ Further, in the nonbridge (OP1/OP2) positions, sulfur will prevent binding of hard divalent metal ions such as Mg^{2+} , which is often seen in RNA enzyme active sites, but will bind softer thiophilic divalent ions such as Cd^{2+} . Hence, thio substitution experiments are often used to study thio/rescue effects in order to probe for functional divalent metal ion binding sites. In the absence of divalent metal ion binding, these positions have much smaller effects that involve a balance between electronic and solvation effects.

Here, we apply ML training procedures using not only the native reaction but also reactions of S2', S5', and S12 variants. As described above, we use a 6 Å range correction and TREMD data at 298, 315, and 330 K as initial data. However, here, we explicitly examine ML parameter sets that were trained both without and with the active learning approach. We focus on the analysis of the results that use DFTB2 as the base QM model and MNDO/d as the target QM model (i.e., DFTB2 + ML \rightarrow MNDO/d). The corresponding results switching the base and target QM models are provided in Supporting Information.

Table 4 quantifies the improvement gained by active learning and the transferability of the trained models. The table summarizes the reaction profiles by listing the reaction free energies (ΔG) and transition-state barriers ΔG^{\ddagger} . The MNDO/d QM/MM profiles are the reference profiles, and the ML corrections are applied to the DFTB2 system. All ML corrections use 1 and 6 Å for R_{on} and R_{off} , respectively, and the initial training set was generated with the TREMD@ 298,315,330 protocol (ML 8 in Table 2). The initial training set includes the energies and forces from all 91 windows spanning the reaction coordinate of the native ligand and the 91 windows for each of the S2', S5', and S12 thio-substituted ligands. The S3' and S1P thio substitution data are not included in the training of the ML corrections. The training to the initial data produces four sets of ML parameters (by running the optimization four times with different random number seeds). DFTB2 + ML free-energy profiles are

generated for each set of ML parameters. The four free-energy profiles are averaged, and the columns in Table 4 labeled "ML (No AL)" are taken from the average profile. Similarly, the columns labeled "ML (with AL)" are taken from an averaged profile after applying nine cycles of AL. Each cycle of AL produced several additional frames per umbrella window simulation, all of which were included in the labeling phases.

Overall, the training procedure appears to be remarkably robust (Figure 3), despite the fact that for the reactions of the variants, the DFTB2 and MNDO/d profiles are strikingly different compared with that of the native reaction. Specifically, the DFTB2 sulfur is considerably softer, leading to dramatically more stable alkyl sulfide species. This results in reaction ΔG values that are much more positive and negative for the S2' (nucleophile position) and S5' (leaving group position) variant reactions, respectively. The reactions all occur through an inline 2' nucleophilic attack to form a pentavalent dianionic phosphorane transition state (or intermediate), followed by departure of the 5' leaving group. In the transition state/ intermediate, the 2' and 5' positions occupy axial positions in the pentavalent phosphorane (X2'-P-X5' angle roughly 180°), and the XP1, XP2, and X3' positions are located in equatorial positions. The X3' position is a "bridging" position between the phosphorus and C3' atom, whereas XP1 and XP2 are terminal "nonbridge" positions. DFTB2 predicts that sulfur in an equatorial position in the phosphorane is greatly overstabilized (by more than 30 kcal/mol), leading to a deep artificially stable phosphorane intermediate not observed for MNDO/d. Nonetheless, for the reactions included in the training (Figure 3a,c,d,f), the DFTB2 + ML models are in remarkable agreement with the MNDO/d profiles. ML training in the absence of AL provides good overall agreement with the shape of the profiles but tends to have barriers that are still underpredicted. In all cases, further training with AL leads to improvement of the barriers. After AL, the greatest error of reactions used in training with respect to the MNDO/d reference free-energy barriers occurs for the native and S12 reactions, with errors of -0.3 and -0.9 kcal/mol, respectively

pubs.acs.org/JCTC

Article



$DFTB2+ML \rightarrow MNDO/d$

Figure 3. Free-energy profiles of different variants listed in Table 1, compared with original DFTB2 and MNDO/d curves. Curves labeled ML (with AL) were calculated from four independent MD simulations at the last cycle of AL. Curves labeled ML (no AL) were calculated from one MD simulation at the first cycle of AL.

(roughly 3% error in the forward barrier). Errors in the reaction free-energy values are all below 1.0 kcal/mol in magnitude.

In order to assess transferability of the DFTB2 + ML model, we examine 1D profiles for two reactions of S1P and S3' (Figure 3b,e) that were not considered in the training. DFTB2 + ML models the S1P reaction reasonably well, whereas for the S3' reaction, the barrier is considerably underestimated. The latter is due to the overstabilization of equatorial sulfur in the DFTB2 model. The S3' variant is the only example in all the reactions where there is a sulfur in a bridging equatorial position in the pentavalent phosphorane transition state (whereas this is a stable intermediate with pure DFTB2). Hence, the neural network did not have data representative of this situation, and the DFTB2 + ML correction is in greater error. The S1P reaction also has a sulfur in an equatorial position in the pentavalent phosphorane transition state; however, it is nonbridging. In the training of DFTB2 + ML, the S12 reaction had two sulfur atoms in nonbridging positions, and this enabled the resulting network to better predict the behavior for the S1P variant which had only one. One striking feature of the data shown in Table 4 is the improvement that the AL procedure affords, which is most pronounced for the

reactions of variants not even considered in the training. After AL, the S1P reaction barriers had an error that was reduced from 1.5 kcal/mol to only 0.5 kcal/mol (2.2% error), whereas the S3' reaction barrier had an error that was reduced from -1.4 to 0.1 kcal/mol (0.5% error). This is strongly suggestive that the AL procedure is a powerful method to create more robust and transferable ML models.

We further examined transferability by considering 2D freeenergy profiles for the native reaction from umbrella sampling simulations not included in the training. Figure 4 compares the 2D free-energy surfaces ($R_{P-O2'}$ and $R_{P-O5'}$ coordinates) and 1D reaction profiles (derived from the 2D surface) for the native reaction using MNDO/d + ML \rightarrow DFTB2 (left panels) and DFTB2 + ML \rightarrow MNDO/d (right panels). Only the active learning ML correction is considered. Transition-state barriers and reaction coordinate values are provided in Supporting Information. Overall, for the native reaction, AL very accurately models the full 2D free-energy surface, reaction path, and transition-state barrier and geometry for both MNDO/d and DFTB2 target reference data.

3.4. Behavior and Computational Performance of the DPRc Potential. In order to illustrate the behavior of the DPRc QM/MM interaction range correction, Figure 5 shows



Figure 4. 2D free-energy surfaces ($R_{P-O2'}$ and $R_{P-O5'}$ coordinates) and 1D reaction profiles (derived from the 2D surface) for the native reaction using MNDO/d + ML \rightarrow DFTB2 (left panels) and DFTB2 + ML \rightarrow MNDO/d (right panels). Only the AL ML correction is shown. The top row shows the 2D surface for target QM reference results (left, DFTB2; right, MNDO/d) as well as a 1D profile along the reaction path obtained from the 2D surface and projected onto the 1D reaction coordinate $\xi_{PT} = R_{P-O5'} - R_{P-O2'}$ (shown as insets, with the transition-state position labeled with a solid dot). The bottom row shows the 2D surface for base QM + ML model results (left, MNDO/d + ML; right, DFTB2 + ML) as well as a 1D profile along the reaction path obtained from the 2D surface and projected onto the 1D reaction coordinate $\xi_{PT} = R_{P-O5'} - R_{P-O2'}$ (shown as insets, with the transition-state position labeled with a solid dot).



Figure 5. Interaction energy ΔE between the native QM solute and a single MM water calculated at a range of phosphorous (P) and water oxygen (Ow) distances by 1D rigid translation of the water along the phosphate plane. The model ML 8 in Table 2 was used to correct the interaction energy from DFTB2 to the MNDO/d level. The purpose of this plot is to simply illustrate the nature of the DPRc term for a simple well-defined interaction energy between QM and MM molecules.

the interaction energy between an MM (TIP4P/Ew) water molecule with the QM solute as a function of the phosphorus– water oxygen (P-OW) interaction distance. Shown are the DFTB2 and MNDO/d QM/MM potential energies and a DFTB2 + ML model trained to reproduce the MNDO/d QM/ MM energies and forces. The ML parameterization is the same used within Table 4. The interaction potential is seen to be quite smooth and quickly approaches 0 by 4.0 Å until it is rigorously and smoothly turned off by 6 Å. In this way, the DPRc term is able to compensate for differences in the shortranged exchange repulsions as well as mid-range electrostatic interactions between the MNDO/d and DFTB2 QM models.

The inclusion of the DPRc correction to the QM/MM calculations necessarily causes the simulations to require more wall clock time to complete the simulation. Table 5 compares the wall clock time of the DFTB QM/MM simulations shown in Figure 2b as the DPRc QM/MM cutoff is increased. The timings are shown when the entirety of the calculation is performed on one CPU core and when the QM/MM calculation is performed on one CPU core and the DPRc correction is accelerated by use of a GPU coprocessor. The

Table 5. Wall Clock Time per Simulation Step (ms/step) Observed in the DFTB2 QM/MM Simulations of the Native Reaction with and without the Use of DPRc Corrections^a

	QM/MM				
		0	3	6	9
CPU	140	374	442	748	1416
GPU		152	152	156	164

"The timings were performed using a 2.10 GHz Intel Xeon Gold 6230 CPU with 28 MB of cache. The "CPU" times performed the QM/MM and DPRc corrections on one CPU core. The "GPU" times performed the QM/MM calculation on one CPU core and the DPRc correction on an NVIDIA V100 GPU. The "QM/MM" column does not include a DPRc correction. The remaining columns include the DPRc model to correct the interaction between QM atoms and the interactions between QM and MM atoms separated by a distance not greater than $R_{\rm C}$, where $R_{\rm C} \equiv R_{\rm off,QM/MM}$ is the cutoff for the smooth range correction). The timings with DPRc correction correspond to models 1b–4b in Table 2.

simulations were performed using a 2.10 GHz Intel Xeon Gold 6230 CPU with 28 MB of cache and a NVIDIA V100 GPU. Figure 6 further illustrates the scaling of the DPRc term as a function of the QM/MM range correction cutoff.



Figure 6. Illustration of scaling of the wall clock time per simulation step (ms/step) observed in the DFTB2 QM/MM simulations of the native reaction with and without the use of DPRc corrections. Times are shown as a function of $R_C \equiv R_{off,QM/MM}$ the cutoff for the smooth range correction for the QM/MM interactions (a value of 0 indicates no range correction). These values are listed in Table 5 and described in more detail in the caption. The DPRc timings could be fit to a nonlinear model $t_0 + a \cdot (R_C)^b$, where t_0 is the timing in ms/step for the DPRc model with $R_C = 0$ Å, *a* is a pre-factor parameter in units of (ms/step)/(Å^b), and *b* is an empirical unitless scaling exponent. The simulations were performed using a 2.10 GHz Intel Xeon Gold 6230 CPU with 28MB of cache and an NVIDIA V100 GPU.

The timings shown in Table 5 suggest that the DPRc correction is very expensive if performed on CPUs. Extending the correction to include 6 Å of the MM solvent is 5 times more expensive than performing an uncorrected DFTB2 QM/ MM force evaluation. Alternatively, the DPRc correction is relatively inexpensive if it is performed on a GPU owing to our use of DeePMD-kit, which has been highly tuned for GPU performance. The GPU evaluation of the DPRc model using a 6 Å correction increases the cost of the underlying QM/MM (single-core CPU) calculation by only 11%. The cost of an *ab initio* QM Hamiltonian is orders of magnitude more expensive than a semiempirical evaluation in comparison, which is the primary motivation for developing the DPRc correction. The cost of the DPRc roughly

pubs.acs.org/JCTC

linearly with number of QM/MM particle–particle interactions, which in the limit of very large values of QM/MM interaction range correction cutoff, $R_{off,QM/MM}$, is expected to be proportional to $(R_{off,QM/MM})^3$ (i.e., proportional to the volume, assuming a uniform density of MM particles around the QM solute). In the present example, the timings for both CPU and CPU/GPUs can be fit to a scaling model as illustrated in Figure 6 (see Supporting Information for further details). Examination of the pre-factors suggests that the evaluation of the range correction term is roughly 100 times faster using an NVIDIA V100 GPU coprocessor than using a single CPU core alone.

3.5. Discussion of the DPRc Model and Training Procedure in the Context of Similar-Spirited Approaches and Future Applications. As discussed above, the ultimate goal of this work is to develop a DPRc training procedure that can be applied to correct a fast semiempirical QM/MM model to the accuracy level of an *ab initio* QM/MM model. We fully acknowledge that the training procedures developed here might require more data sampling or further adaptation to be transferable to ab initio training data. Nonetheless, this underscores the need to develop as robust and reliable a procedure as possible. In the present work, we develop a training procedure using fast, approximate base and target QM models. Our specific choice of MNDO/d¹³ and DFTB2¹⁴ was made for two reasons. First, both methods are sufficiently fast that we can extensively sample both 1D and 2D free-energy profiles for the reactions of interest. This allows us to explore different training procedures and rigorously evaluate them by assessing errors of the final profiles since it is possible to also compute the profiles to high precision at the target QM level. This would not be possible with an *ab initio* QM model. Second, MNDO/d and DFTB2 were not chosen because either method is particularly accurate. Rather, they were chosen because they differ quite substantially in their energetic behavior for both internal QM energies and QM/MM electrostatic interaction energies and as a result give quite different profiles for the reactions being considered. For example, the MNDO approximation is known to underestimate torsion energy barriers due to the lack of an overlap matrix in the generalized eigenvalue problem, which is corrected with DFTB2. This is especially problematic for MNDO/d to obtain the correct puckering of five-membered rings^{87-89} as is needed for the present reactions that contain a five-membered ring in the reactant state and form an additional five-membered ring in the course of the reaction. A specialized AM1/d-PhoT model has been developed,⁸⁴ along with sugar pucker correction terms⁸⁹ to improve this behavior. On the other hand, DFTB2 uses a crude atom-centered monopolar representation for electrostatic QM/MM interactions, whereas MNDO/d uses a full multipolar representation. In this way, MNDO/d and DFTB2 provide models that have considerably different internal QM energies and generate different ensembles as conformational ensembles as well as substantially different mid-range QM/MM interaction energies. Thus, the choice of MNDO/d and DFTB2 represents a considerable stress test for any training procedure. While one cannot conclusively assert that any successful procedure developed here will be completely transferable to training against *ab initio* QM/MM target data, it is the hope that our choice of contrasting base and target QM models will lead to a robust training procedure that is at least largely transferable. It should be noted as well that our strategy moving forward is to depart

from the most accurate fast, approximate QM models and using a quantum mechanical force field (QMFF) strategy^{11,16,91} where even with DFTB-like models, multipolar mapping parameters can be developed to improve the quality of the electrostatic potential for QM/MM interactions to be very close to that of *ab initio* DFT methods.^{8,16}

It is worth mentioning that there are similar procedures being actively pursued in the literature with promising success. $^{28-31,33,34,47-51,55,92}$ These technologies are still in very early stages, so it is premature to attempt to evaluate their relative merits, but nonetheless should be pointed out and acknowledged. Recent works have used ML training to match semiempirical model forces to ab initio reference data to examine the free-energy surface of glycine condensation.^{17,54} Specifically, Kroonblawd et al.¹⁷ used approximately 450 reference configurations obtained from condensed phase simulations to adjust the short-range "repulsive" potentials within the DFTB2 semiempirical QM model. The initial set of reference data used in the present work (1000 configurations per temperature) is approximately twice as large. It is likely that our use of an AL procedure could allow us to initiate the training with a smaller set of data. Another distinction between refs 17 and 54 and the present work is our use of ML corrections to the QM/MM interaction in addition to the QM/QM interactions. Other very recent approaches have used force matching on collective variables in QM/MM simulations using semiempirical models to calculate free-energy surfaces³⁵, and machine-learning assisted models for enzyme reactions in solution.⁵⁰ All these procedures have promise and a great deal of complementarity, and thus, we anticipate the landscape for practical QM/MM simulations to undergo exciting transformations that open doors to new applications.

4. CONCLUSIONS

In this work, we develop a new DPRc ML potential, along with an AL procedure for neural network training. We test the DPRc model and training procedure against a series of six nonenzymatic phosphoryl transfer reactions in solution that are important in mechanistic studies of RNA-cleaving enzymes. We used DPRc model corrections to a base QM model in order to reproduce reference data from a QM target model. For this purpose, we chose the MNDO/d and DFTB2 approximate quantum models that differ from one another substantially in order to provide a rigorous stress test for the DPRc model and training procedure. Examination of different range corrections and training data for the native reaction suggest that range corrections for QM/MM interactions out to 6 Å are required for high accuracy and further that there is a benefit from including data from enhanced sampling with temperature replica exchange, including some data at elevated temperatures to train models to avoid high-energy regions of conformational space. The DPRc model was demonstrated to be highly robust in its ability to accurately model four different reactions simultaneously (i.e., with the same universal model). Further, the DPRc model and training procedure were demonstrated to be highly transferable to model 2D freeenergy surfaces for the native reaction, along with a 1D freeenergy profile for the reaction of the S1P variant, both of which were not explicitly considered in the training. Transferability to the reaction of the S3' variant was not as successful, producing a free-energy profile that was overall correct in shape, but this underestimated the forward barrier by 5.5 kcal/mol. Here, it was shown that AL had a profound effect on improving the

pubs.acs.org/JCTC

DPRc model in terms of agreement with training data as well as transferability outside the scope of the training data. It should be noted that the DP-GEN scheme used to develop the DPRc models considers the atomic forces as target parameters to train the neural networks. Our assessment and validation of the models, however, consider the analyzed free-energy surfaces that result from umbrella sampling simulations along the reaction coordinate(s) which were not explicitly considered in the training per se. Hence, the success of the models reported here has important implications for quantitative freeenergy simulations that use range-corrected ML potentials to enhance the accuracy of fast, approximate QM/MM methods. It is the hope that the new range-corrected ML potential, along with the robust training procedure, will enable the creation of next-generation QM/MM and QMFF potentials for a wide spectrum of free-energy applications, ranging from drug discovery to enzyme design.

ASSOCIATED CONTENT

③ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00201.

Training errors (energy unit: eV, force unit: eV/Å) for the native reaction with different ML parameter sets; ratio of frames with different model deviations in the final simulations for the native reaction with different ML parameter sets; comparison of target (DFTB2) and model (MNDO/d+ML) transition-state coordinate values; transition-state (TS) barrier height and reaction coordinate values for the 2D study of the native reaction, with DFTB2 and MNDO/d Hamiltonians and ML models trained only on 1D reaction; training errors and histogram of frame energies; forces in the QM region; forces in the MM region; free-energy profiles of different variants, compared with original DFTB2 and MNDO/d curves; and illustration of scaling of the wall clock time per simulation step (ms/step) observed in the DFTB2 QM/MM simulations of the native reaction with and without the use of DPRc corrections CPU and GPU timings plotted separately(PDF)

AUTHOR INFORMATION

Corresponding Author

Darrin M. York – Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine, and Department of Chemistry and Chemical Biology, Rutgers the State University of New Jersey, New Brunswick, New Jersey 08901-8554, United States; orcid.org/0000-0002-9193-7055; Email: Darrin.York@rutgers.edu

Authors

- Jinzhe Zeng Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine, and Department of Chemistry and Chemical Biology, Rutgers the State University of New Jersey, New Brunswick, New Jersey 08901-8554, United States; Orcid.org/0000-0002-1515-8172
- **Timothy J. Giese** Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine, and Department of Chemistry and Chemical Biology, Rutgers the State University of New Jersey, New Brunswick, New Jersey

pubs.acs.org/JCTC

08901-8554, United States; © orcid.org/0000-0002-0653-9168

Şölen Ekesan – Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine, and Department of Chemistry and Chemical Biology, Rutgers the State University of New Jersey, New Brunswick, New Jersey 08901-8554, United States; ◎ orcid.org/0000-0002-5598-5754

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c00201

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

J.Z. would like to thank help from Dr. Han Wang and Dr. Linfeng Zhang on development of DeePMD-kit. The authors are grateful for financial support provided by the National Institutes of Health (grant GM107485 to D.M.Y.). Computational resources were provided by the Office of Advanced Research Computing (OARC) at Rutgers, The State University of New Jersey (specifically, the Amarel cluster and associated research computing resources), and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant ACI-1548562 (specifically, the resources COMET at SDSC through allocation TG-CHE190067). The authors also acknowledge the Texas Advanced Computing Center (TACC, http://www.tacc.utexas.edu) at The University of Texas at Austin for providing HPC resources, specifically the Frontera Supercomputer, that have contributed to the research results reported within this paper. J.Z. is grateful for the Excellence Fellowship for the doctoral study provided by Rutgers University.

REFERENCES

(1) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99*, 1129–1143.

(2) Brooks, B. R.; Brooks, C. L., III; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. J. Comput. Chem. **2009**, *30*, 1545–1614.

(3) Lee, T.-S.; Cerutti, D. S.; Mermelstein, D.; Lin, C.; LeGrand, S.; Giese, T. J.; Roitberg, A.; Case, D. A.; Walker, R. C.; York, D. M. GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *J. Chem. Inf. Model.* **2018**, *58*, 2043–2050.

(4) Cui, Q.; Pal, T.; Xie, L. Biomolecular QM/MM Simulations: What Are Some of the "Burning Issues"? *J. Phys. Chem. B* 2021, 125, 689–702.

(5) Ganguly, A.; Weissman, B. P.; Giese, T. J.; Li, N.-S.; Hoshika, S.; Rao, S.; Benner, S. A.; Piccirilli, J. A.; York, D. M. Confluence of theory and experiment reveals the catalytic mechanism of the Varkud satellite ribozyme. *Nat. Chem.* **2020**, *12*, 193–201.

(6) Lee, T.-S.; Allen, B. K.; Giese, T. J.; Guo, Z.; Li, P.; Lin, C.; McGee, T. D., Jr; Pearlman, D. A.; Radak, B. K.; Tao, Y.; Tsai, H.-C.; Xu, H.; Sherman, W.; York, D. M. Alchemical Binding Free Energy Calculations in AMBER20: Advances and Best Practices for Drug Discovery. J. Chem. Inf. Model. 2020, 60, 5595–5623.

(7) Gao, J. Hybrid Quantum and Molecular Mechanical Simulations: An Alternative Avenue to Solvent Effects in Organic Chemistry. *Acc. Chem. Res.* **1996**, *29*, 298–305.

(8) Giese, T. J.; Panteva, M. T.; Chen, H.; York, D. M. Multipolar Ewald methods, 2: Applications using a quantum mechanical force field. *J. Chem. Theory Comput.* **2015**, *11*, 451–461.

(9) Giese, T. J.; York, D. M. Ambient-Potential Composite Ewald Method for ab Initio Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulation. *J. Chem. Theory Comput.* **2016**, *12*, 2611–2632.

(10) Ito, S.; Cui, Q. Multi-level free energy simulation with a staged transformation approach. J. Chem. Phys. **2020**, 153, 044115.

(11) Giese, T. J.; York, D. M. Quantum mechanical force fields for condensed phase molecular simulations. *J. Phys. Condens. Matter* 2017, 29, 383002.

(12) Thiel, W. Semiempirical quantum-chemical methods. Wiley Interdiscip. Rev.: Comput. Mol. Sci. 2014, 4, 145–157.

(13) Thiel, W.; Voityuk, A. A. Extension of MNDO to d orbitals: parameters and results for silicon. *J. Mol. Struct.: THEOCHEM* **1994**, 313, 141–154.

(14) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2001**, *105*, 569–585.

(15) Hourahine, B.; Aradi, B.; Blum, V.; Bonafé, F.; Buccheri, A.; Camacho, C.; Cevallos, C.; Deshaye, M. Y.; Dumitrică, T.; Dominguez, A.; Ehlert, S.; Elstner, M.; van der Heide, T.; Hermann, J.; Irle, S.; Kranz, J. J.; Köhler, C.; Kowalczyk, T.; Kubař, T.; Lee, I. S.; Lutsker, V.; Maurer, R. J.; Min, S. K.; Mitchell, I.; Negre, C.; Niehaus, T. A.; Niklasson, A. M. N.; Page, A. J.; Pecchia, A.; Penazzi, G.; Persson, M. P.; Řezáč, J.; Sánchez, C. G.; Sternberg, M.; Stöhr, M.; Stuckenberg, F.; Tkatchenko, A.; Yu, V. W.-z.; Frauenheim, T. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. J. Chem. Phys. 2020, 152, 124101.

(16) Giese, T. J.; Huang, M.; Chen, H.; York, D. M. Recent Advances toward a General Purpose Linear-Scaling Quantum Force Field. *Acc. Chem. Res.* **2014**, *47*, 2812–2820.

(17) Kroonblawd, M. P.; Pietrucci, F.; Saitta, A. M.; Goldman, N. Generating Converged Accurate Free Energy Surfaces for Chemical Reactions with a Force-Matched Semiempirical Model. *J. Chem. Theory Comput.* **2018**, *14*, 2207–2218.

(18) Mueller, T.; Hernandez, A.; Wang, C. Machine learning for interatomic potential models. J. Chem. Phys. 2020, 152, 050902.

(19) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401–146404.

(20) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **2017**, *8*, 3192–3203.

(21) Smith, J. S.; Roitberg, A. E.; Isayev, O. Transforming Computational Drug Discovery with Machine Learning and AI. ACS Med. Chem. Lett. **2018**, *9*, 1065–1069.

(22) Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching Coupled Cluster Accuracy with a General-purpose Neural Network Potential Through Transfer Learning. *Nat. Commun.* **2019**, *10*, 2903.

(23) Nebgen, B.; Lubbers, N.; Smith, J. S.; Sifain, A. E.; Lokhov, A.; Isayev, O.; Roitberg, A. E.; Barros, K.; Tretiak, S. Transferable Dynamic Molecular Charge Assignment Using Deep Neural Networks. J. Chem. Theory Comput. **2018**, *14*, 4687–4698.

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

(25) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet - A Deep Learning Architecture for Molecules and Materials. *J. Chem. Phys.* **2018**, *148*, 241722.

pubs.acs.org/JCTC

(26) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.* **2019**, *15*, 3678–3693.

(27) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120*, 145301.

(28) Zhang, Y.; Hu, C.; Jiang, B. Embedded Atom Neural Network Potentials: Efficient and Accurate Machine Learning with a Physically Inspired Representation. *J. Phys. Chem. Lett.* **2019**, *10*, 4962–4967.

(29) Bartók, A. P.; Csányi, G. Gaussian Approximation Potentials: A Brief Tutorial Introduction. *Int. J. Quantum Chem.* **2015**, *115*, 1051–1057.

(30) Shao, Y.; Hellström, M.; Mitev, P. D.; Knijff, L.; Zhang, C. PiNN: A Python Library for Building Atomic Neural Networks of Molecules and Materials. *J. Chem. Inf. Model.* **2020**, *60*, 1184–1193.

(31) Khorshidi, A.; Peterson, A. A. Amp: A Modular Approach to Machine Learning in Atomistic Simulations. *Comput. Phys. Commun.* **2016**, 207, 310–324.

(32) Zhang, L.; Han, J.; Wang, H.; Car, R.; Weinan, E. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* **2018**, *120*, 143001.

(33) Chen, Y.; Zhang, L.; Wang, H.; Weinan, E. DeePKS: A Comprehensive Data-Driven Approach toward Chemically Accurate Density Functional Theory. *J. Chem. Theory Comput.* **2021**, *17*, 170–181.

(34) Lindsey, R. K.; Fried, L. E.; Goldman, N.; Bastea, S. Active learning for robust, high-complexity reactive atomistic simulations. *J. Chem. Phys.* **2020**, *153*, 134117.

(35) Gastegger, M.; McSloy, A.; Luya, M.; Schütt, K. T.; Maurer, R. J. A deep neural network for molecular wave functions in quasi-atomic minimal basis representation. *J. Chem. Phys.* **2020**, *153*, 044123.

(36) Zhang, L.; Han, J.; Wang, H.; Saidi, W.; Car, R.; Weinan, E. Advances in Neural Information Processing Systems 31; Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R., Eds.; Curran Associates, Inc., 2018; pp 4436–4446.

(37) Galib, M.; Limmer, D. T. Reactive uptake of N 2 O 5 by atmospheric aerosol is dominated by interfacial processes. *Science* **2021**, 371, 921–925.

(38) Zeng, J.; Cao, L.; Xu, M.; Zhu, T.; Zhang, J. Z. H. Complex reaction processes in combustion unraveled by neural network-based molecular dynamics simulation. *Nat. Commun.* **2020**, *11*, 5713.

(39) Zhang, L.; Lin, D.-Y.; Wang, H.; Car, R.; Weinan, E. Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys. Rev. Mater.* **2019**, *3*, 23804.

(40) Zhang, Y.; Wang, H.; Chen, W.; Zeng, J.; Zhang, L.; Wang, H.; Weinan, E. DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun.* **2020**, *253*, 107206.

(41) Calegari Andrade, M. F.; Ko, H.-Y.; Zhang, L.; Car, R.; Selloni, A. Free energy of proton transfer at the water-TiO2 interface from ab initio deep potential molecular dynamics. *Chem. Sci.* **2020**, *11*, 2335–2341.

(42) Zhang, L.; Chen, M.; Wu, X.; Wang, H.; Weinan, E.; Car, R. Deep neural network for the dielectric response of insulators. *Phys. Rev. B* **2020**, *102*, 41121.

(43) Zeng, J.; Zhang, L.; Wang, H.; Zhu, T. Exploring the Chemical Space of Linear Alkane Pyrolysis via Deep Potential GENerator. *Energy Fuels* **2021**, *35*, 762–769.

(44) Wang, H.; Zhang, L.; Han, J.; Weinan, E. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **2018**, 228, 178–184.

(45) Lu, D.; Wang, H.; Chen, M.; Lin, L.; Car, R.; Weinan, E.; Jia, W.; Zhang, L. 86 PFLOPS Deep Potential Molecular Dynamics simulation of 100 million atoms with ab initio accuracy. *Comput. Phys. Commun.* **2021**, *259*, 107624.

(46) Jia, W.; Wang, H.; Chen, M.; Lu, D.; Lin, L.; Car, R.; Weinan, E.; Zhang, L. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis; SC '20; IEEE Press, 2020.

(47) Shen, L.; Wu, J.; Yang, W. Multiscale Quantum Mechanics/ Molecular Mechanics Simulations with Neural Networks. J. Chem. Theory Comput. **2016**, *12*, 4934–4946.

(48) Wu, J.; Shen, L.; Yang, W. Internal force corrections with machine learning for quantum mechanics/molecular mechanics simulations. *J. Chem. Phys.* **2017**, *147*, 161732.

(49) Shen, L.; Yang, W. Molecular Dynamics Simulations with Quantum Mechanics/Molecular Mechanics and Adaptive Neural Networks. *J. Chem. Theory Comput.* **2018**, *14*, 1442–1455.

(50) Pan, X.; Yang, J.; Van, R.; Epifanovsky, E.; Ho, J.; Huang, J.; Pu, J.; Mei, Y.; Nam, K.; Shao, Y. Machine-Learning-Assisted Free Energy Simulation of Solution-Phase and Enzyme Reactions. *J. Chem. Theory Comput.* **2021**, *17*, 5745–5758.

(51) Böselt, L.; Thürlemann, M.; Riniker, S. Machine Learning in QM/MM Molecular Dynamics Simulations of Condensed-Phase Systems. J. Chem. Theory Comput. **2021**, *17*, 2641–2658.

(52) Young, T.; Johnston-Wood, T.; Deringer, V.; Duarte, F. A Transferable Active-Learning Strategy for Reactive Molecular Force Fields; Royal Society of Chemistry, 2021.

(53) Yue, S.; Muniz, M. C.; Calegari Andrade, M. F.; Zhang, L.; Car, R.; Panagiotopoulos, A. Z. When do short-range atomistic machinelearning models fall short? *J. Chem. Phys.* **2021**, *154*, 034111.

(54) Lindsey, R. K.; Kroonblawd, M. P.; Fried, L. E.; Goldman, N. Force Matching Approaches to Extend Density Functional Theory to Large Time and Length Scales. In *Computational Approaches for Chemistry Under Extreme Conditions*; Goldman, N., Ed.; Springer International Publishing: Cham, 2019; pp 71–93.

(55) Kim, B.; Snyder, R.; Nagaraju, M.; Zhou, Y.; Ojeda-May, P.; Keeton, S.; Hege, M.; Shao, Y.; Pu, J. Reaction Path-Force Matching in Collective Variables: Determining Ab Initio QM/MM Free Energy Profiles by Fitting Mean Force. *J. Chem. Theory Comput.* **2021**, *17*, 4961–4980.

(56) Case, D. A.; Belfon, K.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kovalenko, K.; Krasny, R.; Kurtzman, T.; Lee, T.; Le-Grand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Man, V.; Merz, K.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Skrynnikov, N.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wilson, R. M.; Wolf, R. M.; Wu, X.; Xiong, Y.; Xue, Y.; York, D. M.; Kollman, P. A. *AMBER 20*; University of California: San Francisco: San Francisco, CA, 2020.

(57) Emilsson, G. M.; Nakamura, S.; Roth, A.; Breaker, R. R. Ribozyme speed limits. *RNA* **2003**, *9*, 907–918.

(58) Gaines, C. S.; Giese, T. J.; York, D. M. Cleaning Up Mechanistic Debris Generated by Twister Ribozymes Using Computational RNA Enzymology. ACS Catal. 2019, 9, 5803-5815. (59) Gaines, C. S.; York, D. M. Model for the Functional Active State of the TS Piberyme from Molecular Simulation Argent Charac

State of the TS Ribozyme from Molecular Simulation. Angew. Chem., Int. Ed. 2017, 129, 13577–13580.

(60) Ekesan, Ş.; York, D. M. Dynamical ensemble of the active state and transition state mimic for the RNA-cleaving 8-17 DNAzyme in solution. *Nucleic Acids Res.* **2019**, *47*, 10282–10295.

(61) Kostenbader, K.; York, D. M. Molecular simulations of the pistol ribozyme: unifying the interpretation of experimental data and establishing functional links with the hammerhead ribozyme. *RNA* **2019**, 25, 1439–1456.

(62) Gaines, C. S.; Piccirilli, J. A.; York, D. M. The L-platform/L-scaffold framework: a blueprint for RNA-cleaving nucleic acid enzyme design. *RNA* **2020**, *26*, 111–125.

(63) Giese, T. J.; Ekesan, Ş.; York, D. M. Extension of the Variational Free Energy Profile and Multistate Bennett Acceptance Ratio Methods for High-Dimensional Potential of Mean Force Profile Analysis. *J. Phys. Chem. A* **2021**, *125*, 4216–4232.

(64) Giese, T. J.; York, D. M. Development of a Robust Indirect Approach for MM \rightarrow QM Free Energy Calculations That Combines

pubs.acs.org/JCTC

Force-Matched Reference Potential and Bennett's Acceptance Ratio Methods. J. Chem. Theory Comput. 2019, 15, 5543–5562.

(65) Bevilacqua, P. C.; Harris, M. E.; Piccirilli, J. A.; Gaines, C.; Ganguly, A.; Kostenbader, K.; Ekesan, Ş.; York, D. M. An Ontology for Facilitating Discussion of Catalytic Strategies of RNA-Cleaving Enzymes. *ACS Chem. Biol.* **2019**, *14*, 1068–1076.

(66) Lee, T.-S.; Radak, B. K.; Pabis, A.; York, D. M. A new maximum likelihood approach for free energy profile construction from molecular simulations. *J. Chem. Theory Comput.* **2013**, *9*, 153–164.

(67) Lee, T.-S.; Radak, B. K.; Huang, M.; Wong, K.-Y.; York, D. M. Roadmaps through free energy landscapes calculated using the multidimensional vFEP approach. *J. Chem. Theory Comput.* **2014**, *10*, 24–34.

(68) Giese, T. J.; York, D. M. FE-ToolKit: The Free Energy Analysis Toolkit. https://gitlab.com/RutgersLBSR/fe-toolkit.

(69) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. J. Cheminf. **2011**, *3*, 33.

(70) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(71) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J. Chem. Theory Comput. **2015**, *11*, 3696–3713.

(72) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. J. Comput. Chem. 2000, 21, 132–146.

(73) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. J. Chem. Phys. 2004, 120, 9665–9678.

(74) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N?-methylamide. *Biopolymers* **1992**, *32*, 523–535.

(75) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(76) Nam, K.; Gao, J.; York, D. M. An efficient linear-scaling Ewald method for long-range electrostatic interactions in combined QM/ MM calculations. J. Chem. Theory Comput. 2005, 1, 2–13.

(77) Gregersen, B. A.; Lopez, X.; York, D. M. Hybrid QM/MM study of thio effects in transphosphorylation reactions. J. Am. Chem. Soc. 2003, 125, 7178–7179.

(78) Gregersen, B. A.; Lopez, X.; York, D. M. Hybrid QM/MM Study of Thio Effects in Transphosphorylation Reactions: The Role of Solvation. *J. Am. Chem. Soc.* **2004**, *126*, 7504–7513.

(79) Range, K.; McGrath, M. J.; Lopez, X.; York, D. M. The Structure and Stability of Biological Metaphosphate, Phosphate, and Phosphorane Compounds in the Gas Phase and in Solution. *J. Am. Chem. Soc.* **2004**, *126*, 1654–1665.

(80) Giese, T. J.; Gregersen, B. A.; Liu, Y.; Nam, K.; Mayaan, E.; Moser, A.; Range, K.; Faza, O. N.; Lopez, C. S.; Lera, A. R. d.; Schaftenaar, G.; Lopez, X.; Lee, T.-S.; Karypis, G.; York, D. M. QCRNA 1.0: A database of quantum calculations for RNA catalysis. *J. Mol. Graph. Model.* **2006**, *25*, 423–433.

(81) Liu, Y.; Gregersen, B. A.; Lopez, X.; York, D. M. Density Functional Study of the In-Line Mechanism of Methanolysis of Cyclic Phosphate and Thiophosphate Esters in Solution: Insight into Thio Effects in RNA Transesterification. *J. Phys. Chem. B* **2005**, *109*, 19987–20003.

(82) Weber, W.; Thiel, W. Orthogonalization corrections for semiempirical methods. *Theor. Chem. Acc.* **2000**, *103*, 495–506.

(83) Giese, T. J.; York, D. M. Density-functional expansion methods: grand challenges. *Theor. Chem. Acc.* **2012**, *131*, 1145.

(84) Nam, K.; Cui, Q.; Gao, J.; York, D. M. Specific reaction parametrization of the AM1/d Hamiltonian for phosphoryl transfer

reactions: H, O, and P atoms. J. Chem. Theory Comput. 2007, 3, 486–504.

(85) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method: Third-Order Expansion of the Density Functional Theory Total Energy and Introduction of a Modified Effective Coulomb Interaction. J. Phys. Chem. A 2007, 111, 10861–10873.

(86) Yang, Y.; Yu, H.; York, D.; Elstner, M.; Cui, Q. Description of phosphate hydrolysis reactions with the self-consistent-charge density-functional-tight-binding (SCC-DFTB) theory. 1. Parameterization. J. Chem. Theory Comput. 2008, 4, 2067–2084.

(87) Huang, M.; Giese, T. J.; Lee, T.-S.; York, D. M. Improvement of DNA and RNA Sugar Pucker Profiles from Semiempirical Quantum Methods. J. Chem. Theory Comput. 2014, 10, 1538–1545. (88) Huang, M.; Giese, T. J.; York, D. M. Nucleic acid reactivity: Challenges for next-generation semiempirical quantum models. J. Comput. Chem. 2015, 36, 1370–1389.

(89) Huang, M.; Dissanayake, T.; Kuechler, E.; Radak, B. K.; Lee, T.-S.; Giese, T. J.; York, D. M. A Multidimensional B-Spline Correction for Accurate Modeling Sugar Puckering in QM/MM Simulations. J. Chem. Theory Comput. **2017**, *13*, 3975–3984.

(90) Dantzman, C. L.; Kiessling, L. L. Reactivity of a 2'-Thio Nucleotide Analog. J. Am. Chem. Soc. 1996, 118, 11715–11719.

(91) Giese, T. J.; York, D. M. A Modified Divide-and-Conquer Linear-Scaling Quantum Force Field with Multipolar Charge Densities; Cui, Q., Meuwly, M., Ren, P., Eds.; Pan Stanford, 2016; Chapter 1.

(92) Glick, Z. L.; Metcalf, D. P.; Koutsoukas, A.; Spronk, S. A.; Cheney, D. L.; Sherrill, C. D. AP-Net: An atomic-pairwise neural network for smooth and transferable interaction potentials. *J. Chem. Phys.* **2020**, *153*, 044112.