

RESEARCH ARTICLE | JUNE 10 2024

## Amber free energy tools: Interoperable software for free energy simulations using generalized quantum mechanical/molecular mechanical and machine learning potentials F FREE

Special Collection: [Modular and Interoperable Software for Chemical Physics](#)

Yujun Tao ; Timothy J. Giese ; Şölen Ekesan ; Jinzhe Zeng ; Bálint Aradi ; Ben Hourahine ; Hasan Metin Aktulga ; Andreas W. Götz ; Kenneth M. Merz, Jr. ; Darrin M. York  



*J. Chem. Phys.* 160, 224104 (2024)  
<https://doi.org/10.1063/5.0211276>



The Journal of Chemical Physics

Special Topic:  
Dynamic Exciton for Materials, Biology  
and Energy Conversion

Guest Editors: Hiroshi Imahori, Prashant Kamat, Hironori Kaji, Yasuhiro Kobori, Kiminori Maeda, Michael R. Wasielewski  
JCP Editors: Tianquan (Tim) Lian, Renee Frontiera, Jennifer Ogilvie, Qiang Shi

[Submit Today!](#)



# Amber free energy tools: Interoperable software for free energy simulations using generalized quantum mechanical/molecular mechanical and machine learning potentials

Cite as: J. Chem. Phys. 160, 224104 (2024); doi: 10.1063/5.0211276

Submitted: 29 March 2024 • Accepted: 15 May 2024 •

Published Online: 10 June 2024



View Online



Export Citation



CrossMark

Yujun Tao,<sup>1</sup>  Timothy J. Giese,<sup>1</sup>  Şölen Ekesan,<sup>1</sup>  Jinzhe Zeng,<sup>1</sup>  Bálint Aradi,<sup>2</sup>  Ben Hourahine,<sup>3</sup>   
Hasan Metin Aktulga,<sup>4</sup>  Andreas W. Götz,<sup>5</sup>  Kenneth M. Merz Jr.,<sup>4</sup>  and Darrin M. York<sup>1,a)</sup> 

## AFFILIATIONS

<sup>1</sup>Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854, USA

<sup>2</sup>Bremen Center for Computational Materials Science, University of Bremen, D-28334 Bremen, Germany

<sup>3</sup>SUPA, Department of Physics, University of Strathclyde, Glasgow G4 0NG, United Kingdom

<sup>4</sup>Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, USA

<sup>5</sup>San Diego Supercomputer Center, University of California San Diego, La Jolla, California 92093, USA

**Note:** This paper is part of the JCP Special Topic on Modular and Interoperable Software for Chemical Physics.

<sup>a)</sup>Author to whom correspondence should be addressed: [Darrin.York@rutgers.edu](mailto:Darrin.York@rutgers.edu)

## ABSTRACT

We report the development and testing of new integrated cyberinfrastructure for performing free energy simulations with generalized hybrid quantum mechanical/molecular mechanical (QM/MM) and machine learning potentials (MLPs) in Amber. The Sander molecular dynamics program has been extended to leverage fast, density-functional tight-binding models implemented in the DFTB+ and xTB packages, and an interface to the DeePMD-kit software enables the use of MLPs. The software is integrated through application program interfaces that circumvent the need to perform “system calls” and enable the incorporation of long-range Ewald electrostatics into the external software’s self-consistent field procedure. The infrastructure provides access to QM/MM models that may serve as the foundation for QM/MM- $\Delta$ MLP potentials, which supplement the semiempirical QM/MM model with a MLP correction trained to reproduce *ab initio* QM/MM energies and forces. Efficient optimization of minimum free energy pathways is enabled through a new surface-accelerated finite-temperature string method implemented in the FE-ToolKit package. Furthermore, we interfaced Sander with the i-PI software by implementing the socket communication protocol used in the i-PI client-server model. The new interface with i-PI allows for the treatment of nuclear quantum effects with semiempirical QM/MM- $\Delta$ MLP models. The modular interoperable software is demonstrated on proton transfer reactions in guanine-thymine mispairs in a B-form deoxyribonucleic acid helix. The current work represents a considerable advance in the development of modular software for performing free energy simulations of chemical reactions that are important in a wide range of applications.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0211276>

## INTRODUCTION

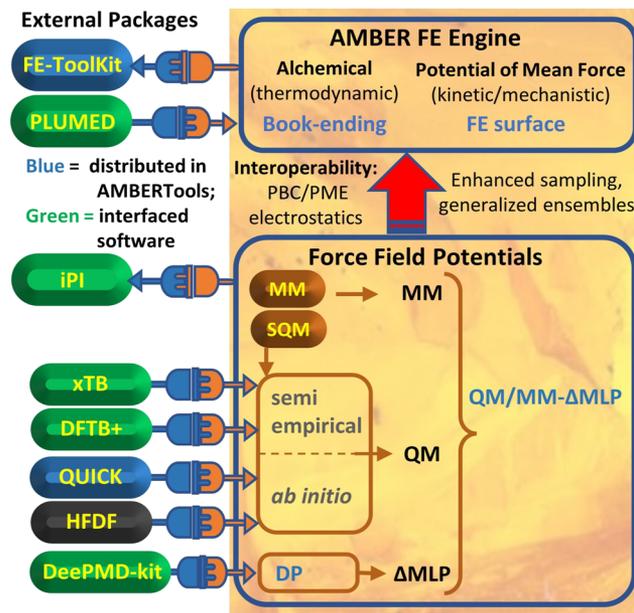
Free energy simulations are used to study a wide range of reactive chemical processes in the condensed phase.<sup>1,2</sup> The calculation of multi-dimensional free energy surfaces<sup>3,4</sup> (FES) and minimum free energy paths<sup>5-9</sup> (MFEP) can be used to predict mechanisms. The

insight gained by these predictions can aid in the interpretation of experimental data, guide new experiments, and ultimately inform the design of new technology. *Ab initio* quantum mechanical models can provide high accuracy in these simulations; however, their computational cost can severely limit the size of the quantum system and/or the degree of sampling that can be realized. This limitation is

further amplified if path integral molecular dynamics (PIMD)<sup>10–15</sup> is required to introduce nuclear quantum effects, as these simulations are even more computationally intensive.

An attractive alternative to *ab initio* quantum mechanical/molecular mechanical (QM/MM) simulation is the design of quantum mechanical force fields<sup>2,16,17</sup> and machine learning models.<sup>18–23</sup> Of particular relevance to the current work is the development of QM/MM- $\Delta$ MPL models, whereby the energies and forces of a fast, approximate QM model are corrected with machine-learning potential.<sup>20,24–30</sup> These models have the potential to offer the computational efficiency needed to address complex chemical mechanisms that require sampling of high-dimensional free energy surfaces while providing accuracy comparable to high level QM methods. A barrier to progress in the development and validation of such methods is their availability in flexible software packages that enable a wide range of applications in the condensed phase.

The purpose of the present work is to report the development and testing of interoperable simulation software in Amber for performing free energy simulations (including PIMD) with QM/MM- $\Delta$ MPL models. A list of software components is itemized in Table I and Fig. 1 illustrates their interoperation. Specifically, we have extended the QM/MM capabilities of the Sander molecular dynamics program by introducing interfaces to the DFTB+<sup>42</sup> and xTB<sup>38–41</sup> semiempirical QM packages. This enables access to a wide range of powerful new density-functional tight-binding models with enabled dispersion and hydrogen bonding corrections for use in condensed phase QM/MM simulations with rigorous long-ranged electrostatics under periodic boundary conditions.<sup>37,52–54</sup> Furthermore, we incorporated the DeePMD-kit software<sup>43–45</sup> into Sander<sup>31</sup> to evaluate new machine-learning potentials. Path integral molecular dynamics (PIMD) has been enabled through a new interface with the i-PI software.<sup>46</sup> Finally, a recently developed surface-accelerated finite-temperature string method<sup>9</sup> has been incorporated into the FE-ToolKit software<sup>51</sup> (distributed within AmberTools<sup>31</sup>) to locate minimum free energy paths.



**FIG. 1.** Interoperable software in Amber for conducting free energy simulations in the condensed phase with QM/MM- $\Delta$ MPL models. An Amber-colored rectangle indicates the Sander program with general internal components encircled by blue. Arrows indicate the calling sequence. Green ovals are external software packages not distributed with Amber/AmberTools, whereas blue ovals indicate new cyber-infrastructure distributed with Amber. The program names and functionality are described in Table 1.

The Methods section describes the interfaces between Sander, the DFTB+ and xTB semiempirical packages, the DeePMD-kit software, and i-PI. The nature of the QM/MM interactions and the incorporation of Ewald electrostatics within the self-consistent field

**TABLE I.** Summary of program acronyms and functionality.

Program	Function
Sander	A molecular dynamics program with support for QM/MM potentials <sup>31</sup>
QUICK	A GPU-accelerated <i>ab initio</i> QM program <sup>32–36</sup>
HFDF	An <i>ab initio</i> QM program optimized for large memory CPU machines <sup>37</sup>
xTB	A semiempirical tight-binding program that implements the GFN1-xTB and GFN2-xTB models <sup>38–41</sup>
DFTB+	A self-consistent charge density functional tight binding program that implements the DFTB2 and DFTB3 methods and various dispersion models <sup>42</sup>
DeePMD-kit	A machine learning library for performing deep potential training and inference based on the TensorFlow platform <sup>43–45</sup>
i-PI	A program for performing path integral molecular dynamics <sup>46</sup>
PLUMED	A library that implements enhanced-sampling algorithms, collective variable definitions, and biasing potentials to calculate free energy surfaces <sup>47,48</sup>
WESTPA	An interoperable, scalable software package for weighted ensemble simulation and analysis <sup>49,50</sup>
FE-ToolKit	A collection of programs used to analyze alchemical free energy simulations and umbrella window sampling <sup>51</sup>

(SCF) procedures are discussed. We demonstrate the features of the software framework through an example application that examines guanine-thymine mispair tautomerization reactions in B-DNA. These mispairs are enabled by the formation of rare tautomer forms and have been implicated in the formation of mutations in deoxyribonucleic acid (DNA) replication and translation.<sup>55,56</sup> The Results and Discussion section begins with a validation of the implementation by demonstrating QM/MM and QM/MM- $\Delta$ MPLP energy conservation in simulations performed in the microcanonical ensemble. We then compare free energy profiles generated from *ab initio*, semiempirical, and  $\Delta$ MPLP-corrected semiempirical methods obtained from QM/MM umbrella sampling. Finally, we perform QM/MM umbrella sampling with PIMD to explore how nuclear quantum effects alter the profiles.

## METHODS

The QM/MM- $\Delta$ MPLP force fields supplement the QM/MM energy with a nonelectrostatic correction,  $E_{ML}(\{\mathbf{r}\})$ , where  $\{\mathbf{r}\}$  is the 3N array of atomic positions,

$$E_{QM/MM-\Delta M L P} = E_{MM}(\{\mathbf{r}\}) + E_{QM}(\{\mathbf{r}\}; \mathbf{P}) + E_{QM/MM}(\{\mathbf{r}\}; \mathbf{P}) + E_{ML}(\{\mathbf{r}\}), \quad (1)$$

where  $E_{MM}$  and  $E_{QM}$  are the MM and QM contributions to the energy, and  $E_{QM/MM}$  is their interaction.  $E_{ML}$  is the machine learning correction (referred to as a machine learning potential correction,  $\Delta$ MPLP).  $E_{QM}$  and  $E_{QM/MM}$  depend on the single particle density matrix,  $\mathbf{P}$ . In the present work, we incorporated the DFTB+ and xTB packages into Sander for calculating  $E_{QM}$ , and the interface to the DeepMD-kit is used to evaluate  $E_{ML}$ . The details of the implementation and a description of the QM/MM interactions are discussed below.

## Sander interface with DFTB+ and xTB

The DFTB+ and xTB source codes<sup>57,58</sup> are not distributed with AmberTools, but they can be downloaded and compiled as standalone applications and/or software libraries to be used in other applications. In our implementation, we directly link the DFTB+ and xTB libraries with the Sander executable during compilation. Unlike many of the other QM/MM interfaces supported by Sander, this strategy circumvents the cost associated with writing an input file for the external QM program, executing the external program via a system call, and reading the external program's output file at each molecular dynamics step. Instead, Sander directly interacts with DFTB+ and xTB via function calls. Furthermore, this allows us to incorporate variational long-range Ewald electrostatics into their SCF procedures.<sup>37,52-54</sup> Detailed instructions for configuring Sander with DFTB+ and xTB support are provided in the AmberTools 2024 user manual.

The QM/MM interaction is performed with electrostatic embedding,

$$E_{QM/MM} = E_{QM/MM,LJ} + \sum'_{b \in MM} \iint \frac{q_{QM}(\mathbf{r}') q_b(\mathbf{r} - \mathbf{r}_b)}{|\mathbf{r} - \mathbf{r}'|} d^3 r d^3 r' + \sum_{a \in QM} q_a \left( \Delta \phi_{MM}(\mathbf{r}_a) + \frac{1}{2} \Delta \phi_{QM}(\mathbf{r}_a) \right) \quad (2)$$

where  $E_{QM/MM,LJ}$  is the Lennard-Jones interaction between the QM and MM atoms. The second term is the electrostatic interaction between the QM charge density (determined from the density matrix),  $q_{QM}(\mathbf{r})$ , and the field of nearby MM charges,  $q_b(\mathbf{r} - \mathbf{r}_b)$ . The primed summation denotes that only those MM atoms within a cutoff of the QM region are included. The last term accounts for the long-range electrostatic interactions. The QM charge density is modeled by Mulliken charges,  $q_a$ , where  $\Delta \phi_{MM}(\mathbf{r}_a)$  is the electrostatic potential caused by all MM atoms outside of the cutoff and  $\Delta \phi_{QM}(\mathbf{r}_a)$  is the interaction between the QM region and its periodic images,

$$\Delta \phi_{MM}(\mathbf{r}_a) = \text{Re} \sum_{\mathbf{k} \neq 0} \frac{4\pi}{k^2 V} e^{i\mathbf{k}^T \cdot \mathbf{r}_a - \frac{k^2}{4\beta^2}} \sum_{b \in MM} q_b e^{-i\mathbf{k}^T \cdot \mathbf{r}_b} - \sum'_{b \in MM} \frac{q_b}{r_{ab}} \frac{\text{erf}(\beta r_{ab})}{r_{ab}} - \frac{\pi Q_{MM}}{\beta^2 V}, \quad (3)$$

$$\Delta \phi_{QM}(\mathbf{r}_a) = \text{Re} \sum_{\mathbf{k} \neq 0} \frac{4\pi}{k^2 V} e^{i\mathbf{k}^T \cdot \mathbf{r}_a - \frac{k^2}{4\beta^2}} \sum_{b \in QM} q_b e^{-i\mathbf{k}^T \cdot \mathbf{r}_b} - \sum_{b \in QM} \frac{q_b}{r_{ab}} \frac{\text{erf}(\beta r_{ab})}{r_{ab}} - \frac{\pi Q_{QM}}{\beta^2 V}. \quad (4)$$

The first term in Eqs. (3) and (4) is the reciprocal space potential of the MM and QM charge densities, respectively.  $V$  is the volume of the unit cell,  $\beta$  is the Ewald coefficient,  $\mathbf{k}$  is the angular wave number of the plane wave, and  $Q_{MM}$  and  $Q_{QM}$  are the net charges of the MM and QM regions. The second term in Eqs. (3) and (4) is a real-space correction that removes the potential caused by the nearby Ewald Gaussian functions. The third term in Eqs. (3) and (4) is a neutralizing uniform background correction for charged systems. The MM energy,  $E_{MM}(\{\mathbf{r}\})$ , contains an analogous background correction,  $-\pi Q_{MM}^2/(2\beta^2 V)$ , such that the net correction for the entire system is  $-\pi(Q_{MM} + Q_{QM})^2/(2\beta^2 V)$ .

The Sander program is responsible for calculating  $E_{MM}$ ,  $E_{MM,LJ}$ ,  $\Delta \phi_{MM}$ , and  $\Delta \phi_{QM}$ . The DFTB+ and xTB libraries are responsible for calculating  $E_{QM}$ ,  $q_a$ , and the second term in Eq. (2). A similar separation of responsibilities occurs when calculating atomic forces. At each step of dynamics, Sander evaluates  $\Delta \phi_{MM}$  with the particle mesh Ewald (PME) method,<sup>59,60</sup> and it precomputes the exponentials appearing in Eq. (4). The MM atoms are imaged around the QM region, and a list of MM atoms within a cutoff of any QM atom is generated. A SCF calculation is requested by providing the current set of QM atomic coordinates and the locations, charges, and "hardness" values of the nearby MM atoms. The xTB software uses the hardness values to represent the nearby embedding charges as diffuse monopoles. For example, in GFN2-xTB, the second term in Eq. (2) is given by Eq. (5), where  $s$  indexes the atomic orbital shells and  $q_s$  is a shell-resolved partial charge,

$$\sum'_{b \in MM} \iint \frac{q_{QM}(\mathbf{r}') q_b(\mathbf{r} - \mathbf{r}_b)}{|\mathbf{r} - \mathbf{r}'|} d^3 r d^3 r' \approx \sum'_{b \in MM} \sum_{a \in QM} \sum_{s \in a} \frac{q_s q_b}{\sqrt{r_{ab}^2 + \left[ \frac{\eta_s + \eta_b}{2} \right]^2}}. \quad (5)$$

In the present work, we assign all MM atoms the hardness of hydrogen. The Sander interface with DFTB+ currently treats the MM

atoms as point charges; however, this can be generalized in a similar fashion in the future.

The interfaces to both QM packages take advantage of object oriented features of the Fortran 2008 programming language to communicate the variational contribution of the long-range electrostatics within the SCF procedure. The DFTB+ interface with Sander is contained in the file `dftbplus_module.F90`, located in the Sander source directory of the AmberTools 2024 package. The interface creates a child class that inherits from the `TQDepExtPotGen` data structure provided by the DFTB+ software.<sup>57</sup> This class was specifically designed by the DFTB+ developers to include charge-dependent external potential corrections. The `xTB` interface with Sander is similarly contained in the file `xtb_module.F90` within the Sander source directory. The `xTB` interface inherits from the `TSolvation` structure defined within the `xTB` software.<sup>58</sup> `xTB` normally uses this class to interact with the QM region with an implicit solvent model. In both interfaces, the child class redefines the parent class methods to evaluate the last term in Eq. (2) and the “shift” in the atom potentials:  $\Delta\phi_{\text{MM}}(\mathbf{r}_a) + \Delta\phi_{\text{QM}}(\mathbf{r}_a)$ .

One enables QM/MM calculations by setting `ifqnt = 1` in the Sander input file, which will cause the options in the `&qmmm` Fortran namelist to be read. Setting the option `qm_theory = "DFTBPLUS"` or `qm_theory = "XTB"` will read the interface-specific options in the `&dftbplus` or `&xtb` namelists, respectively. Additional details are described in the upcoming AmberTools 2024 documentation, due to be released in the spring of 2024.

### Sander interface with DeePMD-kit

The design of the DeePMD-kit interface<sup>43–45</sup> with the Sander program<sup>31</sup> is heavily influenced by the development of QM/MM- $\Delta$ MLP potentials.<sup>20,21,61–64</sup> In the Sander input file, the user makes a selection of atoms to be corrected with a MLP. In the context of evaluating a QM/MM- $\Delta$ MLP, the MLP selection is the set of atoms in the QM region. One must also provide the name of the file containing the DeePMD-kit neural network definition and parameters. The Deep-Potential Range-Corrected model (DPRC)<sup>20</sup> corrects the interactions between the QM atoms and the interactions between the QM and nearby MM atoms in a manner that conserves energy. That is to say, the MLP correction to the QM/MM interactions smoothly approaches zero as the MM atom reaches a specified cutoff distance. When using a DPRC model, the Sander input file must also specify the MLP cutoff for the QM/MM interactions. Previous parametrizations of DPRC-based MLP models have used a 6 Å cutoff, which is also used in the present work. At each step of dynamics, Sander will request the MLP energy and forces from the DeePMD-kit by providing the position and “type” of each QM and (optionally) nearby MM atoms. A QM atom type is its 2-character element symbol, whereas the MM atom type, from the perspective of DeePMD-kit, is the element symbol prefixed with the letter “m.” This strategy allows the neural network to correct a QM/QM interaction differently than a QM/MM interaction,

The DPRC energy is the sum of atomic contributions,  $E_i$ ,

$$E_{\text{DPRC}} = \sum_{i=1}^N E_i(\mathbf{r}_i, \{\mathbf{r}_j\}_{j \in n(i)}), \quad (6)$$

where  $N$  is the number of atoms,  $\mathbf{r}_i$  is the location of atom  $i$ , and  $n(i)$  denotes the set of neighboring atoms (i.e., atoms within the DPRC cutoff radius). The expressions for the  $E_i$  values from the neural network can be found elsewhere.<sup>20,21</sup> The atomic decomposition of the DPRC model is readily amenable to parallel calculation. The DPRC contribution to the energy is activated by setting the `idprc = 1` option within the Sander input `&dprc` Fortran namelist.

As written, the DPRC application programming interface is specific to the DeePMD-kit because it instantiates software objects defined within the DeePMD-kit library. Many emerging MLPs, such as Allegro<sup>65</sup> and Mace,<sup>66</sup> are based on the PyTorch framework.<sup>67</sup> We are currently developing a generic Sander/PyTorch QM/MM- $\Delta$ MLP interface to evaluate these models.

### Sander interface with i-PI

The i-PI software<sup>46</sup> is a standalone molecular dynamics program that supports state-of-the-art path integral sampling,<sup>68</sup> including the PIGLET thermostat, which has been shown to dramatically reduce the expense of computing quantum kinetic energy.<sup>69,70</sup> The path integral dynamics is performed with a ring polymer Hamiltonian, consisting of several replicas (beads) that are harmonically restrained in series. At each time step, the potential energy of each bead must be computed; however, these calculations are independent and can be performed in parallel. The i-PI software defines a client-server communication protocol to evaluate the potential energy of external driver programs. We have implemented the “internet socket” variation of the protocol, where multiple instances of Sander can be launched on different computers. When launching a Sander instance, one must provide the hostname (internet protocol address) and port number used for interprocess communication. These quantities are specified as command-line arguments `-host` and `-port`, respectively. Furthermore, the Sander input file must set the option `imin = 7`, which causes the program to wait for communication from the server process.

### Free energy profiles of tautomer reactions from classical molecular dynamics

We prepared a B-form DNA system (PDB ID: 113D)<sup>71</sup> to demonstrate the functionality of the new software infrastructure. The B-DNA crystal structure<sup>71</sup> has drawn interest due to the presence of two G-T wobble pair mismatches,<sup>55,72,73</sup> which could give rise to errors in replication<sup>74,75</sup> and translation.<sup>76</sup> This subsequently led to the computational investigation of G-T tautomerization reactions with QM/MM.<sup>56</sup>

In the present work, we explore tautomerization reactions involving residues G4 and T21. The B-DNA system consists of 762 solute atoms solvated by 5151 waters in a truncated octahedron with 59.3 Å real-space lattice vector lengths. A total of 13  $\text{Cl}^-$  and 35  $\text{Na}^+$  ions were added to neutralize the charge and achieve a 0.14M ion concentration. The system was prepared using an elaborate 15-step procedure described in Ref. 77 using a MM potential. The preparation procedure involves geometry optimization, heating, solvent annealing, and equilibration of the system density while restraining solute heavy atoms. Over the course of the equilibration procedure, the solute restraints are gradually reduced. In total, this procedure involved 6.2 and 2.6 ns of simulation in the *NVT* and

*NPT* ensembles, respectively, using a 1 fs time step. The system was then equilibrated for an additional 100 ns in the *NPT* ensemble. The temperature was maintained at 298 K with a Langevin thermostat<sup>78</sup> (5.0 ps<sup>-1</sup> collision frequency), and the density was equilibrated at 1 atm using the Berendsen isotropic barostat. The MM potential modeled the DNA with the OL15 force field,<sup>79</sup> and the solvent environment was modeled with SPC/Fw waters<sup>80</sup> and the Li *et al.* ion parameters.<sup>81</sup> Electrostatics were calculated with the particle mesh Ewald method<sup>59,60</sup> using 10 Å real-space cutoffs, a 1 Å reciprocal space grid spacing, and tin-foil boundary conditions.

Free energy profiles of the tautomeric reactions depicted in Fig. 3 were calculated from QM/MM umbrella sampling using a variety of *ab initio* and semiempirical methods. The reactions are decomposed into three steps connecting 3 tautomeric forms. The “wGT” state corresponds to the G-T wobble base pair in which the T21:N3 position is protonated. The “GT\*” state is a tautomer in which the T21:H3 proton is covalently bonded to the T21:O4 position. Similarly, the “G\*T” state is a tautomer in which the H3 proton is bound to the G4:O6. The three steps involve the proton displacement between these states: wGT → GT\*, GT\* → G\*T, and G\*T → wGT. The QM/MM methods compared in this work are PBE0/6-31G\*, GFN2-xTB,<sup>41</sup> DFTB3-D3 DFTB3-D4(2b), DFTB3-D4(3b), and AM1/d.<sup>82-84</sup> The DFTB3-D3 model applies Grimme’s D3 dispersion correction<sup>85,86</sup> to the DFTB3/3ob Hamiltonian.<sup>87,88</sup> Similarly, the DFTB3-D4(2b) and DFTB3-D4(3b) models supplement the DFTB3/3ob method with the D4 dispersion model.<sup>89</sup> The DFTB3-D4(3b) method includes a 3-body contribution to the dispersion, whereas the DFTB3-D4(2b) is limited to 2-body contributions. Furthermore, we parametrized a reaction-specific DPRc correction for the AM1/d model (described in the next section), which will be denoted AM1/d+ML. The QM region consists of 51 atoms (the nucleobase and sugar of G4 and T21), and it has a net neutral charge. The minimum free energy path of each step was independently optimized for every QM/MM model. The optimizations were performed with the surface-accelerated finite temperature string method<sup>9</sup> in the space of 5 reaction coordinates. The reaction coordinates are distance differences meant to represent the transfer of H3 proton and the relative displacement of the hydrogen bond pattern:  $\xi_1 = R_{N3-H3} - R_{O6-H3}$ ,  $\xi_2 = R_{O6-H3} - R_{O4-H3}$ ,  $\xi_3 = R_{N1-H1} - R_{N3-H1}$ ,  $\xi_4 = R_{N1-O2} - R_{N2-O2}$ , and  $\xi_5 = R_{N2-O2} - R_{N1-N3}$ . All umbrella sampling described in this work used 200 kcal mol<sup>-1</sup> Å<sup>-2</sup> force constants for each reaction coordinate. The string method began with an initial guess that linearly interpolated the reaction coordinates between the step’s reactant and product states. Each string was sampled with 32 umbrella windows. The initial configurations of each window were sequentially prepared, starting from the reactant state. The QM/MM structure of each window was optimized for 1000 steps of conjugate gradient minimization. The temperature was gradually raised from 0 to 298 K over the course of 40 ps, and an additional 10 ps of QM/MM umbrella sampling was performed in the *NVT* ensemble. 50 iterations of the string method were performed with 4 ps/window/iteration of sampling and a 1 fs time step. Production sampling of the final path was performed for 25 ps, and each production simulation was repeated 4 times with different thermostat random number seeds. The reaction coordinate values were recorded every 10 fs, and the aggregate sampling from all three reaction steps was analyzed with the

multistate Bennett’s acceptance ratio (MBAR) method<sup>90</sup> implemented in the ndfes software<sup>45,51</sup> to produce 5-dimensional free energy surfaces. The profiles presented in the paper are the free energies along the minimum free energy path. The PBE0/6-31G\* simulations are very expensive in relation to the other models. In this case, only 10 iterations of the string method were performed, starting from the optimized AM1/d path. Each string was sampled for 2 ps/window/iteration, and 4 sets of production simulations were run for 10 ps/window/trial with a 1 fs time step. The *ab initio* QM/MM electrostatics were calculated with the ambient-potential composite Ewald method<sup>37</sup> using a 10 Å real-space cutoff, 1 Å reciprocal space grid spacing, and tin-foil boundary conditions.

### Free energy profiles of tautomer reactions from path integral molecular dynamics

We recalculated the free energy profiles from path integral molecular dynamics (PIMD) umbrella sampling performed with the new interface to the i-PI software.<sup>68</sup> The profiles generated from classical and path integral dynamics are compared to explore how nuclear quantum effects change free energy predictions. The PIMD dynamical motion was propagated with 6 beads (replicas) at a 0.25 fs timestep at 298 K using the PIGLET quantum thermostat,<sup>69,91</sup> therefore, up to 6 Sander instances can be launched. The parameters for the PIGLET thermostat were taken from the GLE4MD website.<sup>92,93</sup> The parameters were chosen to reproduce the quantum fluctuations at 298 K and span a range of frequencies up to 4142 cm<sup>-1</sup>. Each step’s path was sampled with 32 umbrella windows using 200 kcal mol<sup>-1</sup> Å<sup>-2</sup> force constants. The PIMD restraint potentials were applied to the centroid positions (the bead-averaged positions) by making use of the i-PI interface to PLUMED.<sup>47,48</sup> PIMD umbrella sampling was performed at 32 windows positioned along the path previously optimized for classical sampling. Each window was sampled for 2.5 ps and repeated 4 times with different thermostat random number seeds. The reaction coordinate values were calculated from the centroid positions, and the free energy surface was generated from MBAR analysis of the aggregate sampling obtained from the 3 reaction steps.

### QM/MM-ΔMLP model training

We demonstrate the new interface between Sander and DeepMD-kit by parametrizing a DPRc correction for the AM1/d-PhoT Hamiltonian. The architecture of the DPRc neural network was previously described in Ref. 94, and we have summarized the model in the [supplementary material](#) for completeness. The AM1/d+ML neural network parameters were optimized with the DP-GEN software<sup>95</sup> to reproduce PBE0/6-31G\* QM/MM energies and forces in the B-DNA system. The DPRc MLP corrects the interactions between the QM atoms and the QM/MM interactions within a 6 Å cutoff. Four sets of neural network parameters were trained against 93 617 samples extracted from the AM1/d-PhoT string calculations using the Adam stochastic optimizer with different random number seeds.<sup>96</sup> The parameter sets were then updated through 7 iterations of active learning. An active learning iteration consists of three steps: exploration, labeling, and retraining. The exploration step uses the current set of neural network parameters to perform QM/MM+ΔMLP umbrella sampling. For each saved sample, the energy and forces produced by the 4 parameter sets are

compared; if the models disagree, then the sample is categorized as a “candidate.” The “labeled” samples are a random selection of (up to) 2000 samples drawn from the pool of candidates. The energy and forces of labeled samples are recalculated with PBE0/6-31G\* QM/MM and included as additional reference data within the next round of training. The active learning procedure terminates when 10% (or fewer) of the samples are selected for labeling.

Each active learning iteration sampled 1920 umbrella windows for 2 ps and saved 50 samples/simulation. These simulations correspond to the first 20 and 40 iterations of the finite temperature string method applied to the  $wGT \rightarrow GT^*$  and  $GT^* \rightarrow G^*T$  reaction steps, respectively. A sample was considered a candidate if the standard deviation in the MLP energy corrections was larger than  $10^{-4}$  eV/atom or the standard deviation in the force correction exceeded 0.08 eV/Å for any atom. The active learning procedure terminated after 7 iterations, requiring a total of 9293 *ab initio* QM/MM evaluations. Neural network parameters were produced from  $4 \times 10^6$  steps of Adam stochastic optimization with a learning rate that exponentially decays from  $10^{-3}$  to  $5 \times 10^{-8}$ .<sup>96</sup> As previously described, production sampling of the B-DNA tautomeric reactions was simulated 4 times with different random numbers of seeds. The AM1/d+ML production sampling was repeated with each of the 4 neural network parameter sets (1 trial/parameter set), and the results were averaged.

## RESULTS AND DISCUSSION

### Validation tests: Force evaluation and energy conservation in NVE simulations

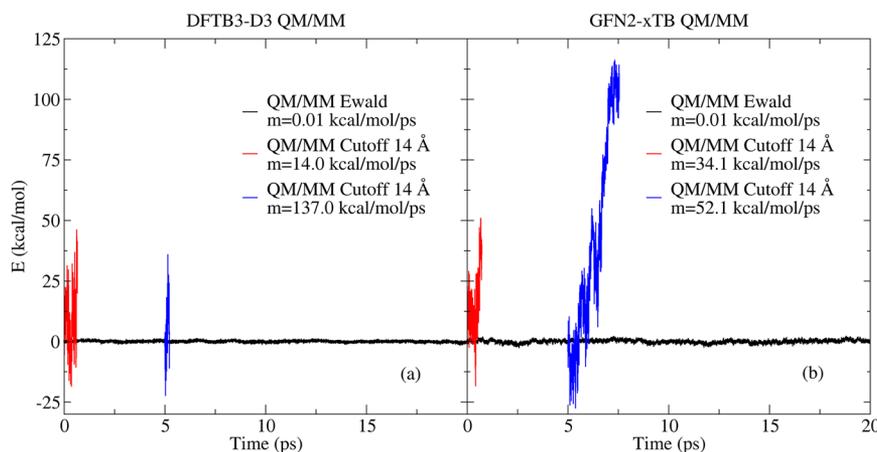
In this section, we validate the Sander interface with DFTB+ and xTB software by performing DFTB3-D3 and GFN2-xTB QM/MM simulations of the B-DNA system in the microcanonical ensemble and monitor the total energy. Energy conservation

is achieved when the potential energy is smooth and the analytic atomic forces are numerically consistent with the potential energy gradients.

Figure 2 plots the total energy of the B-DNA reactant state as a function of time. The lines labeled “QM/MM Ewald” were simulated with Ewald electrostatics using 10 Å real-space cutoffs, whereas “QM/MM Cutoff 14 Å” truncates the QM/MM electrostatics at 14 Å (and it does not model the long-range QM/MM electrostatics). We ran the simulations with truncated electrostatics twice, starting from the GFN2-xTB (Ewald) or DFTB3-D3 (Ewald) configurations at times 0 and 5 ps. The “m” values are the slopes of the energy from linear regression. Each simulation was performed for 20 ps using a 1 fs time step.

The QM/MM Ewald interface described in the methods section assumes that the short-range QM/MM electrostatic interactions [the second term in Eq. (2)] behave like point monopoles at the real-space cutoff; the Ewald correction models the electrostatics outside of the cutoff from a point charge approximation. The short-range electrostatics [Eq. (5)] adopt the Mataga–Nihshimoto–Ohno–Klopman<sup>97–99</sup> damped Coulomb interaction model. This model mimics a point charge interaction at large distances.

Figure 2 suggests that the slight discrepancy between Eq. (5) and point charge electrostatics is negligible when using a 10 Å real-space cutoff for both GFN2-xTB and DFTB3-D3. The energy conservation of GFN2-xTB is almost identical to DFTB3-D3 in DFTB+. Both models exhibit negligible energy drift over 20 ps (slopes of 0.01 kcal/mol/ps). However, when the GFN2-xTB and DFTB3-D3 QM/MM electrostatics are truncated at 14 Å, large deviations in the total energy are encountered, and the self-consistent field procedure often fails to converge within a few picoseconds of simulation. Similar behavior has been illustrated elsewhere in validation tests that compare the stability of QM/MM simulations with and without long-ranged electrostatic interactions.<sup>37</sup>



**FIG. 2.** QM/MM energy conservation of B-DNA in the microcanonical ensemble. (a) DFTB3-D3 performed with Sander/DFTB+. (b) GFN2-xTB performed with Sander/xTB. All methods compute the MM/MM electrostatics with particle mesh Ewald. The QM/MM Ewald method similarly models the QM/QM and QM/MM electrostatics with Ewald. The QM/MM cutoff 14 Å method truncates the QM/MM interactions beyond 14 Å, and long-range QM/QM interactions are ignored. The QM/MM cutoff 14 Å simulations were repeated at different starting configurations taken from the QM/MM Ewald trajectory. The “m” values are the slope of the energy obtained from linear regression.

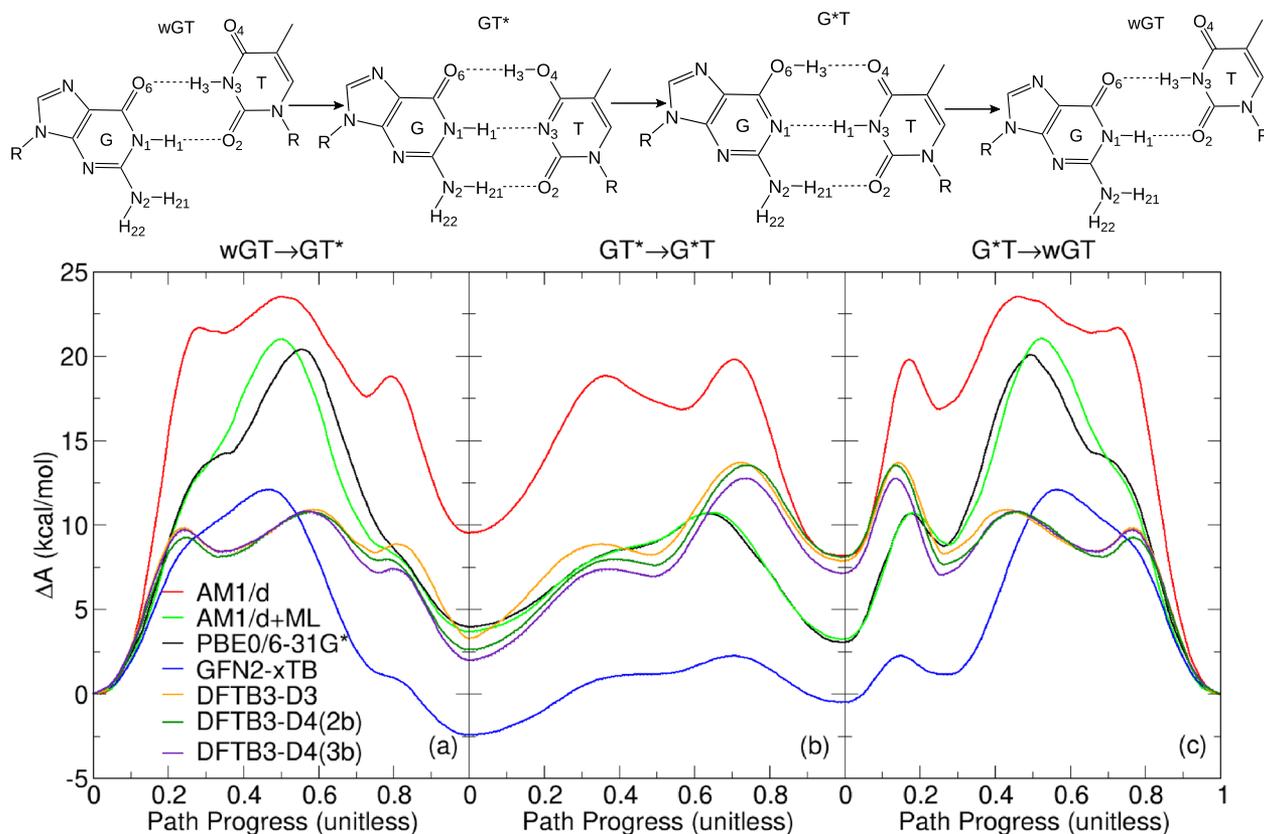
### Free energy surfaces for guanine-thymine mispair tautomerization reactions

In this section, we examine the free energy profiles of guanine-thymine mispair tautomerization reactions using a wide range of *ab initio* and semiempirical/DFTB QM/MM and QM/MM- $\Delta$ MPL models. Free energy surfaces (FESs) and minimum free energy paths (MFEPs) were determined using the surface-accelerated string method<sup>9</sup> with five different reaction coordinates, as described in the Methods section. Figure 3 compares the 1D free energy profiles for the MFEPs for the following QM/MM models: PBE0/6-31G\*, AM1/d, GFN2-xTB, DFTB3-D3, DFTB3-D4(2b), and DFTB3-D4(3b). PBE0/6-31G\* is an *ab initio* QM/MM model, and it is ~500 times slower than the other semiempirical/DFTB QM/MM models for these systems. Several QM/MM models based on DFTB3 are available through the interface with DFTB+, and in particular, the use of different dispersion models is examined. The notation “D3” and “D4” refers to the dispersion correction; (2b) and (3b) mean 2-body and 3-body interaction, respectively, as described in the Methods section. In addition, the GFN2-xTB model is examined through an interface with the xTB package. The AM1/d model is an NDDO-based semiempirical Hamiltonian available within the

SQM module of Sander. AM1/d also serves as the base QM model upon which the QM/MM- $\Delta$ MPL was developed (AM1/d + ML) through an interface with DeePMD-kit.<sup>43–45</sup> Neural network parameters were trained to reproduce PBE0/6-31G\* QM/MM energies and forces. In this way, functionality for a wide range of QM/MM models is demonstrated, and their results are compared.

The reaction itself is decomposed and presented in three steps that form a closed thermodynamic cycle, as illustrated at the top of Fig. 3. In the first step, a G-T wobble pair (wGT) undergoes a coupled proton transfer/hydrogen bond register shift to a G-T\* base pair, where T\* represents the thymine enol tautomer. In the second step of the reaction, the GT\* undergoes a second proton transfer/hydrogen bond register shift to form a (G\*)-T base pair, where G\* is the guanine enol tautomer. The third step of the reaction involves a coupled proton transfer/hydrogen bond register shift to return to the original GT wobble base pair.

It is immediately evident from Fig. 3 that there are large variations in the free energy profiles for the various QM/MM models. Most notably, none of the semiempirical/DFTB-based QM/MM models closely reproduce the *ab initio* PBE0/6-31G\* profile. The AM1/d model systematically overestimates the profile with respect



**FIG. 3.** Helmholtz free energy profiles for guanine-thymine mispair tautomerization reactions using different *ab initio* and semiempirical/DFTB QM/MM and QM/MM- $\Delta$ MPL models performed in the canonical ensemble. (a) wGT  $\rightarrow$  GT\*; (b) GT\*  $\rightarrow$  G\*T; (c) G\*T  $\rightarrow$  wGT. “wGT” indicates a G-T wobble pair, and G\* and T\* are non-standard tautomer states of G and T, respectively, as indicated in the figure and described in the text.

to the wGT state, whereas the DFTB models underestimate the profiles with respect to the wGT state. The GFN2-xTB model, in this example, further predicts that the tautomer base pairs (GT\* and G\*T) are more favorable than the GT wobble (wGT), where the nucleobases are in their most stable tautomer keto forms.

Table II (top) lists the values of key stationary points along the free energy profiles calculated with classical molecular dynamics shown in Fig. 3 and compares them to experimentally derived reference values. Overall, the *ab initio* PBE0/6-31G\* QM/MM model does the best job at reproducing the experimentally derived values. The differences in  $\Delta A$  are less than 1 kcal/mol, and the differences in  $\Delta A^\ddagger$  are less than 3.5 kcal/mol. It should be noted that the comparison with the experiment here is only approximate in the sense that the  $\Delta A^\ddagger$  is only crudely approximated from the experimental rates using transition state theory with unit transmission coefficient. Taking the PBE0/6-31G\* values as a high-level theoretical reference, it is clear that the other QM/MM models differ considerably. For example, the difference between PBE0/6-31G\* and the semiempirical/DFTB model wGT  $\rightarrow$  GT\*  $\Delta A^\ddagger$  values ranges from  $-6.4$  to  $5.6$  kcal/mol. Only the machine-learning potential corrected model AM1/d+ML gives results that are close to the *ab initio* values. The  $\Delta A$  and  $\Delta A^\ddagger$  values agree with PBE0/6-31G\* to within 0.5 and 0.8 kcal/mol, respectively. The AM1/d+ML QM/MM model is thus by far the closest agreement with the *ab initio* QM/MM model, despite the uncorrected AM1/d model having very large differences (maximum differences for  $\Delta A$  and  $\Delta A^\ddagger$  of 5.6 and 3.5 kcal/mol for  $\Delta A$  and  $\Delta A^\ddagger$ , respectively).

The main point of this section is to demonstrate that (1) *ab initio* QM/MM simulations provide reasonable agreement with experimental results such that they can provide, at very high computational cost, a valuable molecular-level interpretation and can potentially be used for prediction; (2) existing semiempirical/DFTB models may not be able to accurately describe a chemical process that was outside the scope of their development; and (3) machine-learning potential corrections can provide an avenue toward improving efficient semiempirical/DFTB QM/MM models for specific applications. The software infrastructure reported here enables new types of QM/MM- $\Delta$ MPLP models to be developed and applied in Amber to efficiently compute free energy surfaces and minimum free energy paths.

### Nuclear quantum effects on guanine-thymine mispair tautomerization reactions

In this section, we demonstrate the new interface with the i-PI software by calculating free energy profiles from PIMD simulation, as described in the Methods section. These calculations are considerably more computationally intensive (by roughly an order of magnitude, or more depending on the number of replicas used in the ring polymer Hamiltonian). Nonetheless, inclusion of nuclear quantum effects is important, especially for modeling chemical reactions that involve proton transfer events such as those presented here. Other recent works have addressed the issue of computational cost

TABLE II. Tautomer free energies and forward reaction barriers (kcal/mol) determined from classical and path integral molecular dynamics umbrella sampling.<sup>a</sup>

Method	wGT $\rightarrow$ GT*		GT* $\rightarrow$ G*T		G*T $\rightarrow$ wGT	
	$\Delta A$	$\Delta A^\ddagger$	$\Delta A$	$\Delta A^\ddagger$	$\Delta A$	$\Delta A^\ddagger$
Expt.	4.43	16.88	-0.62	9.21	-3.82	...
Classical MD						
PBE0/6-31G*	3.97(0.04)	20.41(0.05)	-0.91(0.03)	6.72(0.05)	-3.06(0.04)	17.02(0.05)
AM1/d	9.54(0.04)	23.51(0.06)	-1.36(0.03)	10.26(0.03)	-8.16(0.04)	15.36(0.04)
AM1/d+ML	3.67(0.07)	21.01(0.07)	-0.43(0.07)	7.06(0.07)	-3.24(0.07)	17.78(0.07)
GFN2-xTB	-2.44(0.06)	12.09(0.05)	1.94(0.08)	4.68(0.06)	0.50(0.06)	12.59(0.07)
DFTB3-D3	3.27(0.07)	10.91(0.08)	4.58(0.08)	10.41(0.08)	-7.85(0.09)	5.83(0.09)
DFTB3-D4(3b)	2.01(0.10)	10.81(0.09)	5.13(0.12)	10.75(0.09)	-7.14(0.12)	5.59(0.11)
DFTB3-D4(2b)	2.63(0.07)	10.75(0.07)	5.46(0.07)	10.91(0.06)	-8.09(0.07)	5.43(0.06)
PIMD						
AM1/d	7.98(0.07)	21.87(0.08)	-1.46(0.08)	6.90(0.06)	-6.51(0.09)	15.41(0.08)
AM1/d+ML	3.48(0.03)	20.65(0.04)	0.26(0.03)	4.21(0.08)	-3.75(0.04)	16.88(0.04)
GFN2-xTB	-3.10(0.06)	10.93(0.06)	1.70(0.05)	2.47(0.06)	1.34(0.06)	12.26(0.05)
DFTB3-D3	1.68(0.09)	10.59(0.08)	4.41(0.09)	7.96(0.08)	-6.03(0.09)	4.57(0.08)
DFTB3-D4(3b)	-0.57(0.08)	9.99(0.10)	5.12(0.07)	8.12(0.07)	-4.55(0.07)	5.44(0.09)
DFTB3-D4(2b)	2.17(0.07)	9.78(0.08)	4.20(0.08)	7.04(0.07)	-6.37(0.08)	3.40(0.08)

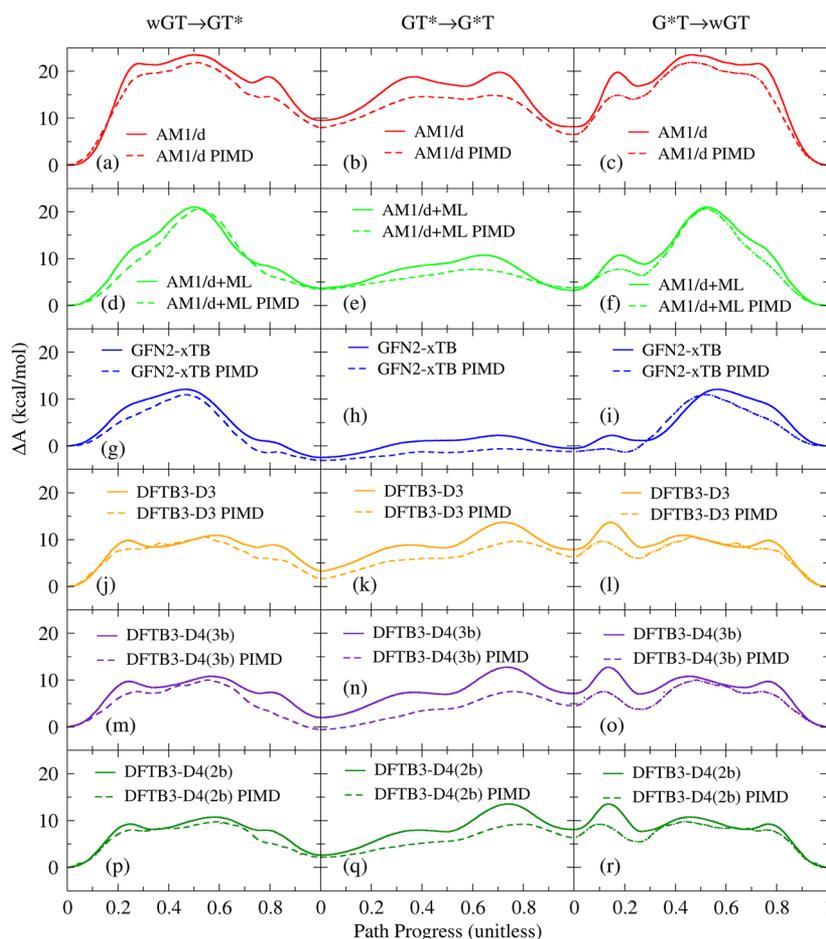
<sup>a</sup>Each free energy profile was computed 4 times from independent sets of simulations. The tabulated values are the means of the 4 trials. The numbers in parentheses are the standard error of the mean. Experimental reference values were derived from rates measured in Ref. 55, and free energy values were approximated from transition state theory assuming a unit transmission coefficient.

in PIMD simulations using fast density-functional tight-binding models<sup>100</sup> or machine learning potentials<sup>10</sup> optimized for high-level reference data.

Figure 4 compares the classical (solid lines) and path-integral molecular dynamics (dashed lines) free energy profiles for different semiempirical/DFTB QM/MM and QM/MM- $\Delta$ MPL models. Table II (bottom) lists the values of key stationary points along the free energy profiles calculated with the PIMD shown in Fig. 4 and compares them to corresponding classical and experimentally derived reference values. Table S2 summarizes the differences between classical and PIMD free energy values. Overall, the inclusion of nuclear quantum effects generally lowers the free energy barrier ( $\Delta A^\ddagger$ ) for proton transfer steps, in some cases exceeding 3 kcal/mol, whereas the effect on  $\Delta A$  is variable (ranging up to roughly  $\pm 2.5$  kcal/mol). The PIMD free energy profiles were calculated from the ring polymer centroid positions. These profiles

generally have lower barriers than those produced from classical molecular dynamics. The apparent smoothing of the profile is due to the delocalization of the nuclear wave packet, which can partially be attributed to tunneling effects.<sup>101–103</sup> The current demonstration examines differences between independently calculated free energy profiles; however, other applications of i-PI can directly evaluate equilibrium and kinetic isotope effects from umbrella sampling using thermodynamic perturbation and frequency factors.<sup>21,104</sup>

The main point of this section is to demonstrate that the inclusion of nuclear quantum effects can have important quantitative effects on the free energy profiles and  $\Delta A^\ddagger$  values for the tautomerization reactions considered here. These calculations are made practical in Amber with fast semiempirical/DFTB QM/MM and QM/MM- $\Delta$ MPL models and enabled by the interface with i-PI.<sup>46</sup>



**FIG. 4.** Classical and path-integral molecular dynamics free energy profiles for guanine-thymine mispair tautomerization reactions using different semiempirical/DFTB QM/MM and QM/MM- $\Delta$ MPL models described in the text: (a) wGT  $\rightarrow$  GT\* (a), (d), (g), (j), (m), (p); (b) GT\*  $\rightarrow$  G\*T (b), (e), (h), (k), (n), (q); (c) G\*T  $\rightarrow$  wGT (c), (f), (i), (l), (o), (r). Here “wGT” indicates a G-T wobble pair, and G\* and T\* are non-standard tautomer states of G and T, respectively. Profiles from classical and PIMD are shown as solid and dashed lines, respectively. The color schemes for different models are the same as those shown in Fig. 3.

**TABLE III.** Classical molecular dynamics performance using Sander with several QM/MM potentials. The values are time steps/day.

Potential	QM/MM Ewald Linked	QM/MM cutoff Linked	QM/MM cutoff File-based
DFTB3-D3	163 000	99 000	80 000
GFN2-xTB	133 000	88 000	80 000
AM1/d	133 000	86 000	...
AM1/d+ML	66 000	51 000	...
PBE0/6-31G*	200	...	...

## Performance

We measured the performance of the B-DNA microcanonical QM/MM simulations shown in Fig. 2 on a single Intel Xeon E5-2630 v3 CPU core clocked at 2.40 GHz. Table III expresses the performance as the number of classical molecular dynamics time steps per day. For example, 1000 steps/day would correspond to 1 ps/day of sampling using a 1 fs time step. More efficient sampling with a longer integration time step may be obtained by using the “middle” thermostat scheme described in Refs. 105 and 106 in conjunction with hydrogen mass repartitioning.<sup>107</sup> The “QM/MM Ewald” timings use 10 Å real-space cutoffs, whereas the “QM/MM Cutoff” timings truncate the QM/MM interactions at 14 Å. The subheadings “Linked” and “File-based” refer to the communication between Sander and the QM program: Linked means there is direct communication via function calls, whereas File-based relies on disk input/output and system calls.

Table III suggests that linking the semiempirical programs into the Sander executable increases the performance by 10%-to-20% over file-based communication. Direct linkage further improves the performance by enabling the use of the QM/MM Ewald method, which achieves stable dynamics with smaller real-space cutoffs. The single core performance of AM1/d, GFN2-xTB, and DFTB3-D3 is very similar. The inclusion of the DPRc  $\Delta$ MMLP halves the AM1/d performance; however, previous examination of DPRc timings found that it decreases the performance by only 10% when the  $\Delta$ MMLP is evaluated on an Nvidia V100 GPU. The AM1/d+ML performance on a single CPU core is 300 times faster than the PBE0/6-31G\* QM/MM evaluation using the HFDF *ab initio* program.

**TABLE IV.** AM1/d QM/MM path integral molecular dynamics performance using i-PI and multiple Sander instances.  $N$  is the number of Sander instances. The path integral dynamics is propagated with 6 beads.

$N$	Steps/day
1	17 000
2	27 000
3	37 000
4	37 000
5	37 000
6	60 000

The high cost of *ab initio* QM/MM evaluation often prevents its use within PIMD simulations. To illustrate the cost of PIMD simulations, we measured the performance of the wGT reactant state discussed in the previous section. The PIMD simulations were run with 6 beads; therefore, a pool of up to 6 Sander instances can be launched. Table IV reports the simulation performance as the number of PIMD time steps evaluated per day. The total number of CPU cores is  $N + 1$ , where  $N$  is the number of Sander instances; the additional core is reserved for the i-PI dynamics program. At each time step, 6 potential energy evaluations are required for the corresponding set of 6 beads; therefore, running 4 or 5 Sander instances does not improve the 3-instance performance because they are not divisors of the number of beads. When 6 instances are used, the performance is about half of what is obtained during classical dynamics due to the communication between the i-PI and the Sander instances.

## CONCLUSION

We reported a new integrated software infrastructure for conducting free energy simulations in the condensed phase under periodic boundary conditions with long-ranged PME/Ewald electrostatics using a wide range of fast QM/MM potentials made available through interfaces with DFTB+ and xTB. Furthermore, integration with the DeePMD-kit enables new QM/MM- $\Delta$ MMLP models to be developed and efficiently applied. Nuclear quantum effects (beyond the harmonic approximation) can be treated using free energy surface sampling with path integral molecular dynamics. The software infrastructure reported here extends the capabilities of Amber for developing new QM/MM- $\Delta$ MMLP models and using them to compute free energy surfaces and related properties for a wide range of applications. Near-future directions for development will involve expanding the scope of interoperability to include interfaces with enhanced sampling software, and in particular with WESTPA,<sup>49,50</sup> a scalable software package for high-performance weighted ensemble simulation and analysis.

## SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for a description of the DPRc neural network and tables of tautomeric reaction free energy differences.

## ACKNOWLEDGMENTS

The authors are grateful for the financial support provided by the National Science Foundation (CSSI Frameworks Grant Nos. 2209717 and 2209718). This work used the supercomputer Frontera<sup>108</sup> at the Texas Advanced Computing Center (TACC) through allocation Grant No. CHE20002, and the supercomputer Expanse at the San Diego Supercomputer Center (SDSC) through allocation Grant No. CHE190067 from the Extreme Science and Engineering Discovery Environment (XSEDE),<sup>109</sup> which was supported by National Science Foundation Grant No. 1548562, and from the Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS) program,<sup>110</sup> which is supported

by National Science Foundation Grant Nos. 2138259, 2138286, 2138307, 2137603, and 2138296.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Yujun Tao:** Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Timothy J. Giese:** Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Şölen Ekesan:** Methodology (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Jinzhe Zeng:** Methodology (equal); Software (equal); Writing – review & editing (equal). **Bálint Aradi:** Writing – review & editing (equal). **Ben Hourahine:** Writing – review & editing (equal). **Hasan Metin Aktulga:** Writing – review & editing (equal). **Andreas Götz:** Writing – review & editing (equal). **Kenneth M. Merz, Jr.:** Writing – review & editing (equal). **Darrin M. York:** Resources (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

### DATA AVAILABILITY

Example Sander input files and the trained DPRc neural network parameter sets for the B-DNA system can be downloaded from [https://gitlab.com/RutgersLBSR/LBSR\\_SuppInfo](https://gitlab.com/RutgersLBSR/LBSR_SuppInfo). We are preparing tutorials describing how to use the new software infrastructure and making them available at <https://gitlab.com/RutgersLBSR/AmberTutorials>.

### REFERENCES

- <sup>1</sup>Free Energy Calculations: Theory and Applications in Chemistry and Biology, edited by C. Chipot and A. Pohorille, Springer Series in Chemical Physics (Springer, New York, 2007), Vol. 86.
- <sup>2</sup>T. J. Giese and D. M. York, “Quantum mechanical force fields for condensed phase molecular simulations,” *J. Phys. Condens. Matter* **29**, 383002 (2017).
- <sup>3</sup>T.-S. Lee, B. K. Radak, M. Huang, K.-Y. Wong, and D. M. York, “Roadmaps through free energy landscapes calculated using the multidimensional vFEP approach,” *J. Chem. Theory Comput.* **10**, 24–34 (2014).
- <sup>4</sup>T. J. Giese, Ş. Ekesan, and D. M. York, “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* **125**, 4216–4232 (2021).
- <sup>5</sup>A. van der Vaart and M. Karplus, “Minimum free energy pathways and free energy profiles for conformational transitions based on atomistic molecular dynamics simulations,” *J. Chem. Phys.* **126**, 164106 (2007).
- <sup>6</sup>H. Hu, Z. Lu, J. M. Parks, S. K. Burger, and W. Yang, “Quantum mechanics/molecular mechanics minimum free-energy path for accurate reaction energetics in solution and enzymes: Sequential sampling and optimization on the potential of mean force surface,” *J. Chem. Phys.* **128**, 34105 (2008).
- <sup>7</sup>E. Vanden-Eijnden and M. Venturoli, “Revisiting the finite temperature string method for the calculation of reaction tubes and free energies,” *J. Chem. Phys.* **130**, 194103 (2009).
- <sup>8</sup>K. Zinovjev and I. Tuñón, “Adaptive finite temperature string method in collective variables,” *J. Phys. Chem. A* **121**, 9764–9772 (2017).
- <sup>9</sup>T. J. Giese, Ş. Ekesan, E. McCarthy, Y. Tao, and D. M. York, “Surface-accelerated string method for locating minimum free energy paths,” *J. Chem. Theory Comput.* **20**, 2058–2073 (2024).
- <sup>10</sup>C. Li and G. A. Voth, “Using machine learning to greatly accelerate path integral *ab initio* molecular dynamics,” *J. Chem. Theory Comput.* **18**, 599–604 (2022).
- <sup>11</sup>J. Gao, “Chapter fourteen—Enzymatic kinetic isotope effects from path-integral free energy perturbation theory,” *Methods Enzymol.* **577**, 359–388 (2016).
- <sup>12</sup>A. Vardi-Kilshtain, N. Nitoker, and D. T. Major, “Nuclear quantum effects and kinetic isotope effects in enzyme reactions,” *Arch. Biochem. Biophys.* **582**, 18–27 (2015).
- <sup>13</sup>J. K. Hwang, Z. T. Chu, A. Yadav, and A. Warshel, “Simulations of quantum mechanical corrections for rate constants of hydride-transfer reactions in enzymes and solutions,” *J. Phys. Chem.* **95**, 8445–8448 (1991).
- <sup>14</sup>J.-K. Hwang and A. Warshel, “A quantized classical path approach for calculations of quantum mechanical rate constants,” *J. Phys. Chem.* **97**, 10053–10058 (1993).
- <sup>15</sup>J.-K. Hwang and A. Warshel, “How important are quantum mechanical nuclear motions in enzyme catalysis?,” *J. Am. Chem. Soc.* **118**, 11745–11751 (1996).
- <sup>16</sup>T. J. Giese, M. Huang, H. Chen, and D. M. York, “Recent advances toward a general purpose linear-scaling quantum force field,” *Acc. Chem. Res.* **47**, 2812–2820 (2014).
- <sup>17</sup>J. Gao, D. G. Truhlar, Y. Wang, M. J. Mazack, P. Löffler, M. R. Provorse, and P. Rehak, “Explicit polarization: A quantum mechanical framework for developing next generation force fields,” *Acc. Chem. Res.* **47**, 2837–2845 (2014).
- <sup>18</sup>J. Behler, “First principles neural network potentials for reactive simulations of large molecular and condensed systems,” *Angew. Chem., Int. Ed.* **56**, 12828–12840 (2017).
- <sup>19</sup>M. Meuwly, “Machine learning for chemical reactions,” *Chem. Rev.* **121**, 10218–10239 (2021).
- <sup>20</sup>J. Zeng, T. J. Giese, Ş. Ekesan, and D. M. York, “Development of range-corrected deep learning potentials for fast, accurate quantum mechanical/molecular mechanical simulations of chemical reactions in solution,” *J. Chem. Theory Comput.* **17**, 6993–7009 (2021).
- <sup>21</sup>T. J. Giese, J. Zeng, Ş. Ekesan, and D. M. York, “Combined QM/MM, machine learning path integral approach to compute free energy profiles and kinetic isotope effects in RNA cleavage reactions,” *J. Chem. Theory Comput.* **18**, 4304–4317 (2022).
- <sup>22</sup>X. Pan, J. Yang, R. Van, E. Epifanovsky, J. Ho, J. Huang, J. Pu, Y. Mei, K. Nam, and Y. Shao, “Machine-learning-assisted free energy simulation of solution-phase and enzyme reactions,” *J. Chem. Theory Comput.* **17**, 5745–5758 (2021).
- <sup>23</sup>R. Snyder, B. Kim, X. Pan, Y. Shao, and J. Pu, “Bridging semiempirical and *ab initio* QM/MM potentials by Gaussian process regression and its sparse variants for free energy simulation,” *J. Chem. Phys.* **159**, 054107 (2023).
- <sup>24</sup>J. Zeng, Y. Tao, T. J. Giese, and D. M. York, “QDπ: A quantum deep potential interaction model for drug discovery,” *J. Chem. Theory Comput.* **19**, 1261–1275 (2023).
- <sup>25</sup>J. Zeng, Y. Tao, T. J. Giese, and D. M. York, “Modern semiempirical electronic structure methods and machine learning potentials for drug discovery: Conformers, tautomers, and protonation states,” *J. Chem. Phys.* **158**, 124110 (2023).
- <sup>26</sup>Y. Han, Z. Wang, Z. Wei, J. Liu, and J. Li, “Machine learning builds full-QM precision protein force fields in seconds,” *Briefings Bioinf.* **22**, 1–10 (2021).
- <sup>27</sup>Y.-J. Zhang, A. Khorshidi, G. Kastlunger, and A. A. Peterson, “The potential for machine learning in hybrid QM/MM calculations,” *J. Chem. Phys.* **148**, 241740 (2018).
- <sup>28</sup>L. Bösel, M. Thürlmann, and S. Riniker, “Machine learning in QM/MM molecular dynamics simulations of condensed-phase systems,” *J. Chem. Theory Comput.* **17**, 2641–2658 (2021).
- <sup>29</sup>L. Shen, J. Wu, and W. Yang, “Multiscale quantum mechanics/molecular mechanics simulations with neural networks,” *J. Chem. Theory Comput.* **12**, 4934–4946 (2016).
- <sup>30</sup>L. Shen and W. Yang, “Molecular dynamics simulations with quantum mechanics/molecular mechanics and adaptive neural networks,” *J. Chem. Theory Comput.* **14**, 1442–1455 (2018).

- <sup>31</sup>D. A. Case, H. M. Aktulga, K. Belfon, D. S. Cerutti, G. A. Cisneros, V. W. D. Cruzeiro, N. Forouzesh, T. J. Giese, A. W. Götz, H. Gohlke, S. Izadi, K. Kasavajhala, M. C. Kaymak, E. King, T. Kurtzman, T.-S. Lee, P. Li, J. Liu, T. Luchko, R. Luo, M. Manathunga, M. R. Machado, H. M. Nguyen, K. A. O'Hearn, A. V. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, A. Risheh, S. Schott-Verdugo, A. Shajan, J. Swails, J. Wang, H. Wei, X. Wu, Y. Wu, S. Zhang, S. Zhao, Q. Zhu, T. E. Cheatham, D. R. Roe, A. Roitberg, C. Simmerling, D. M. York, M. C. Nagan, K. M. Merz, and K. M. AmberTools, *J. Chem. Inf. Model.* **63**, 6183–6191 (2023).
- <sup>32</sup>V. W. D. Cruzeiro, M. Manathunga, K. M. Merz, and A. W. Götz, "Open-source multi-GPU-accelerated QM/MM simulations with AMBER and QUICK," *J. Chem. Inf. Model.* **61**, 2109–2115 (2021).
- <sup>33</sup>Y. Miao and K. M. Merz, "Acceleration of high angular momentum electron repulsion integrals and integral derivatives on graphics processing units," *J. Chem. Theory Comput.* **11**, 1449–1462 (2015).
- <sup>34</sup>M. Manathunga, Y. Miao, D. Mu, A. W. Götz, and K. M. Merz, "Parallel implementation of density functional theory methods in the quantum interaction computational kernel program," *J. Chem. Theory Comput.* **16**, 4315–4326 (2020).
- <sup>35</sup>M. Manathunga, C. Jin, V. W. D. Cruzeiro, Y. Miao, D. Mu, K. Arumugam, K. Keipert, H. M. Aktulga, K. M. Merz, and A. W. Götz, "Harnessing the power of multi-GPU acceleration into the quantum interaction computational kernel program," *J. Chem. Theory Comput.* **17**, 3955–3966 (2021).
- <sup>36</sup>M. Manathunga, H. M. Aktulga, A. W. Götz, and K. M. Merz, "Quantum mechanics/molecular mechanics simulations on NVIDIA and AMD graphics processing units," *J. Chem. Inf. Model.* **63**, 711–717 (2023).
- <sup>37</sup>T. J. Giese and D. M. York, "Ambient-potential composite Ewald method for *ab initio* quantum mechanical/molecular mechanical molecular dynamics simulation," *J. Chem. Theory Comput.* **12**, 2611–2632 (2016).
- <sup>38</sup>C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, and S. Grimme, "Extended tight-binding quantum chemistry methods," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **11**, e01493 (2020).
- <sup>39</sup>S. Grimme, C. Bannwarth, and P. Shushkov, "A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements ( $Z = 1-86$ )," *J. Chem. Theory Comput.* **13**, 1989–2009 (2017).
- <sup>40</sup>S. Grimme, C. Bannwarth, E. Caldeweyher, J. Pisarek, A. Hansen, "A general intermolecular force field based on tight-binding quantum chemical calculations," *J. Chem. Phys.* **147**, 161708 (2017).
- <sup>41</sup>C. Bannwarth, S. Ehlert, and S. Grimme, "GFN2- $\chi$ TB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions," *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).
- <sup>42</sup>B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M. Y. Deshayé, T. Dumitrică, A. Dominguez, S. Ehlert, M. Elstner, T. van der Heide, J. Hermann, S. Irle, J. J. Kranz, C. Köhler, T. Kowalczyk, T. Kubar, I. S. Lee, V. Lutsker, R. J. Maurer, S. K. Min, I. Mitchell, C. Negre, T. A. Niehaus, A. M. N. Niklasson, A. J. Page, A. Pecchia, G. Penazzi, M. P. Persson, J. Rezac, C. G. Sanchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V. W. z. Yu, and T. Frauenheim, "DFTB+, a software package for efficient approximate density functional theory based atomistic simulations," *J. Chem. Phys.* **152**, 124101 (2020).
- <sup>43</sup>H. Wang, L. Zhang, J. Han, and W. E, "DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics," *Comput. Phys. Commun.* **228**, 178–184 (2018).
- <sup>44</sup>W. Liang, J. Zeng, D. M. York, L. Zhang, and H. Wang, in *A Practical Guide to Recent Advances in Multiscale Modeling and Simulation of Biomolecules*, edited by Y. Wang and R. Zhou (AIP Publishing, 2023) Chap. VI, pp 6–1–6–20.
- <sup>45</sup>J. Zeng, D. Zhang, D. Lu, P. Mo, Z. Li, Y. Chen, M. Rynik, L. Huang, Z. Li, S. Shi, Y. Wang, H. Ye, P. Tuo, J. Yang, Y. Ding, Y. Li, D. Tisi, Q. Zeng, H. Bao, Y. Xia, J. Huang, K. Muraoka, Y. Wang, J. Chang, F. Yuan, S. L. Bore, C. Cai, Y. Lin, B. Wang, J. Xu, J.-X. Zhu, C. Luo, Y. Zhang, R. E. A. Goodall, W. Liang, A. K. Singh, S. Yao, J. Zhang, R. Wentzcovitch, J. Han, J. Liu, W. Jia, D. M. York, W. E, R. Car, L. Zhang, and H. Wang, "DeePMD-kit v2: A software package for deep potential models," *J. Chem. Phys.* **159**, 054801 (2023).
- <sup>46</sup>V. Kapil, M. Rossi, O. Marsalek, R. Petraglia, Y. Litman, T. Spura, B. Cheng, A. Cuzzocrea, R. H. Meißner, D. M. Wilkins, B. A. Helfrecht, P. Juda, S. P. Bienvenue, W. Fang, J. Kessler, I. Poltavsky, S. Vandenbrande, J. Wieme, C. Corminboeuf, T. D. Kühne, D. E. Manolopoulos, T. E. Markland, J. O. Richardson, A. Tkatchenko, G. A. Tribello, V. Van Speybroeck, and M. Ceriotti, "1-PI 2.0: A universal force engine for advanced molecular simulations," *Comput. Phys. Commun.* **236**, 214–223 (2019).
- <sup>47</sup>M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. Broglia, and M. Parrinello, "PLUMED: A portable plugin for free-energy calculations with molecular dynamics," *Comput. Phys. Commun.* **180**, 1961–1972 (2009).
- <sup>48</sup>G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, "PLUMED 2: New feathers for an old bird," *Comput. Phys. Commun.* **185**, 604–613 (2014).
- <sup>49</sup>M. C. Zwier, J. L. Adelman, J. W. Kaus, A. J. Pratt, K. F. Wong, N. B. Rego, E. Suárez, S. Lettieri, D. W. Wang, M. Grabe, D. M. Zuckerman, and L. T. Chong, "WESTPA: An interoperable, highly scalable software package for weighted ensemble simulation and analysis," *J. Chem. Theory Comput.* **11**, 800–809 (2015).
- <sup>50</sup>J. D. Russo, S. Zhang, J. M. G. Leung, A. T. Bogetti, J. P. Thompson, A. J. DeGrave, P. A. Torrillo, A. J. Pratt, K. F. Wong, J. Xia, J. Copperman, J. L. Adelman, M. C. Zwier, D. N. LeBard, D. M. Zuckerman, and L. T. Chong, "WESTPA 2.0: High-performance upgrades for weighted ensemble simulations and analysis of longer-timescale applications," *J. Chem. Theory Comput.* **18**, 638–649 (2022).
- <sup>51</sup>T. J. Giese and D. M. York, "FE-Toolkit: The free energy analysis toolkit," <https://gitlab.com/RutgersLSBR/fe-toolkit> (Accessed 3 March 2024).
- <sup>52</sup>K. Nam, J. Gao, and D. M. York, "An efficient linear-scaling Ewald method for long-range electrostatic interactions in combined QM/MM calculations," *J. Chem. Theory Comput.* **1**, 2–13 (2005).
- <sup>53</sup>T. J. Giese, M. T. Panteva, H. Chen, and D. M. York, "Multipolar Ewald methods, 1: Theory, accuracy, and performance," *J. Chem. Theory Comput.* **11**, 436–450 (2015).
- <sup>54</sup>T. J. Giese, M. T. Panteva, H. Chen, and D. M. York, "Multipolar Ewald methods, 2: Applications using a quantum mechanical force field," *J. Chem. Theory Comput.* **11**, 451–461 (2015).
- <sup>55</sup>I. J. Kimsey, E. S. Szymanski, W. J. Zahurancik, A. Shakya, Y. Xue, C.-C. Chu, B. J. Sathyamoorthy, Z. Suo, and H. M. Al-Hashimi, "Dynamic basis for dG•dT misincorporation via tautomerization and ionization," *Nature* **554**, 195–201 (2018).
- <sup>56</sup>P. Li, A. Rangadurai, H. M. Al-Hashimi, and S. Hammes-Schiffer, "Environmental effects on guanine-thymine mispair tautomerization explored with quantum mechanical/molecular mechanical free energy simulations," *J. Am. Chem. Soc.* **142**, 11183–11191 (2020).
- <sup>57</sup>B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M. Y. Deshayé, T. Dumitrică, A. Dominguez, S. Ehlert, M. Elstner, T. van der Heide, J. Hermann, S. Irle, J. J. Kranz, C. Köhler, T. Kowalczyk, T. Kubar, I. S. Lee, V. Lutsker, R. J. Maurer, S. K. Min, I. Mitchell, C. Negre, T. A. Niehaus, A. M. N. Niklasson, A. J. Page, A. Pecchia, G. Penazzi, M. P. Persson, J. Řezáč, C. G. Sánchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V. W.-z. Yu, and T. Frauenheim, "DFTB+: General package for performing fast atomistic calculations," <https://github.com/dftbplus/dftbplus> (Accessed 3 March 2024).
- <sup>58</sup>C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, and S. Grimme, "Semiempirical extended tight-binding program package," <https://github.com/grimme-lab/xtb> (Accessed 3 March 2024).
- <sup>59</sup>T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An  $N \log(N)$  method for Ewald sums in large systems," *J. Chem. Phys.* **98**, 10089–10092 (1993).
- <sup>60</sup>U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh Ewald method," *J. Chem. Phys.* **103**, 8577–8593 (1995).
- <sup>61</sup>P. O. Dral, T. Zubatiuk, and B.-X. Xue, in *Quantum Chemistry in the Age of Machine Learning*, edited by P. O. Dral (Elsevier, 2022), Chap. XXI, pp. 491–507.
- <sup>62</sup>Y. Liu and J. Li, "Permutation-invariant-polynomial neural-network-based  $\Delta$ -machine learning approach: A case for the HO<sub>2</sub> self-reaction and its dynamics study," *J. Phys. Chem. Lett.* **13**, 4729–4738 (2022).
- <sup>63</sup>M. Ruth, D. Gerbig, and P. R. Schreiner, "Machine learning of coupled cluster (T)-energy corrections via delta ( $\Delta$ )-learning," *J. Chem. Theory Comput.* **18**, 4846–4855 (2022).
- <sup>64</sup>P. A. Unzueta, C. S. Greenwell, and G. J. O. Beran, "Predicting density functional theory-quality nuclear magnetic resonance chemical shifts via  $\Delta$ -machine learning," *J. Chem. Theory Comput.* **17**, 826–840 (2021).

- <sup>65</sup>A. Musaelian, S. Batzner, A. Johansson, L. Sun, C. J. Owen, M. Kornbluth, and B. Kozinsky, "Learning local equivariant representations for large-scale atomistic dynamics," *Nat. Commun.* **14**, 579 (2023).
- <sup>66</sup>I. Batatia, D. P. Kovacs, G. N. C. Simm, C. Ortner, and G. Csanyi, "MACE: Higher order equivariant message passing neural networks for fast and accurate force fields," (2022); <https://openreview.net/forum?id=YPPngE-ZU>
- <sup>67</sup>J. Ansel, E. Yang, H. He, N. Gimelshein, A. Jain, M. Voznesensky, B. Bao, P. Bell, D. Berard, E. Burovski, G. Chauhan, A. Chourdia, W. Constable, A. Desmairson, Z. DeVito, E. Ellison, W. Feng, J. Gong, M. Gschwind, B. Hirsh, S. Huang, K. Kalambarkar, L. Kirsch, M. Lazos, M. Lezcano, Y. Liang, J. Liang, Y. Lu, C. K. Luk, B. Maher, Y. Pan, C. Puhersch, M. Reso, M. Saroufim, M. Y. Siraichi, H. Suk, S. Zhang, M. Suo, P. Tillet, X. Zhao, E. Wang, K. Zhou, R. Zou, X. Wang, A. Mathews, W. Wen, G. Chanan, P. Wu, and S. Chintala, "PyTorch 2: Faster machine learning through dynamic python bytecode transformation and graph compilation," in *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems* (Association for Computing Machinery, 2024), Vol. 2.
- <sup>68</sup>M. Ceriotti, J. More, and D. E. Manolopoulos, "i-PI: A Python interface for ab initio path integral molecular dynamics simulations," *Comput. Phys. Commun.* **185**, 1019–1026 (2014).
- <sup>69</sup>M. Ceriotti and D. E. Manolopoulos, "Efficient first-principles calculation of the quantum kinetic energy and momentum distribution of nuclei," *Phys. Rev. Lett.* **109**, 100604 (2012).
- <sup>70</sup>M. Ceriotti and T. E. Markland, "Efficient methods and practical guidelines for simulating isotope effects," *J. Chem. Phys.* **138**, 014112 (2013).
- <sup>71</sup>W. N. Hunter, T. Brown, G. Kneale, N. N. Anand, D. Rabinovich, and O. Kennard, "The structure of guanosine-thymidine mismatches in B-DNA at 2.5 angstroms resolution," *J. Biol. Chem.* **262**, 9962–9970 (1993).
- <sup>72</sup>I. J. Kimsey, K. Petzold, B. Sathyamoorthy, Z. W. Stein, and H. M. Al-Hashimi, "Visualizing transient Watson–Crick-like mispairs in DNA and RNA duplexes," *Nature* **519**, 315–320 (2015).
- <sup>73</sup>E. S. Szymanski, I. J. Kimsey, and H. M. Al-Hashimi, "Direct NMR evidence that transient tautomeric and anionic states in dG•dT form Watson–Crick-like base pairs," *J. Am. Chem. Soc.* **139**, 4326–4329 (2017).
- <sup>74</sup>W. Wang, H. W. Hellinga, and L. S. Beese, "Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis," *Proc. Natl. Acad. Sci. U. S. A.* **108**, 17644–17648 (2011).
- <sup>75</sup>K. Bebenek, L. C. Pedersen, and T. A. Kunkel, "Replication infidelity via a mismatch with Watson–Crick geometry," *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1862–1867 (2011).
- <sup>76</sup>A. Rozov, N. Demeshkina, E. Westhof, M. Yusupov, and G. Yusupova, "New structural insights into translational miscoding," *Trends Biochem. Sci.* **41**, 798–814 (2016).
- <sup>77</sup>Ş. Ekesan and D. M. York, "Dynamical ensemble of the active state and transition state mimic for the RNA-cleaving 8–17 DNAzyme in solution," *Nucleic Acids Res.* **47**, 10282–10295 (2019).
- <sup>78</sup>R. J. Loncharich, B. R. Brooks, and R. W. Pastor, "Langevin dynamics of peptides: The frictional dependence of isomerization rates of *N*-acetylalanine-*N'*-methylamide," *Biopolymers* **32**, 523–535 (1992).
- <sup>79</sup>R. Galindo-Murillo, J. C. Robertson, M. Zgarbova, J. Sponer, M. Otyepka, P. Jurecka, and T. E. Cheatham, "Assessing the current state of AMBER force field modifications for DNA," *J. Chem. Theory Comput.* **12**, 4114–4127 (2016).
- <sup>80</sup>Y. Wu, H. L. Tepper, and G. A. Voth, "Flexible simple point-charge water model with improved liquid-state properties," *J. Chem. Phys.* **124**, 024503 (2006).
- <sup>81</sup>P. Li, B. P. Roberts, D. K. Chakravorty, and K. M. Merz, Jr., "Rational design of particle mesh Ewald compatible Lennard-Jones parameters for +2 metal cations in explicit solvent," *J. Chem. Theory Comput.* **9**, 2733–2748 (2013).
- <sup>82</sup>M. J. S. Dewar, E. Zebisch, E. F. Healy, and J. J. P. Stewart, "Development and use of quantum mechanical molecular models. 76. AM1: A new general purpose quantum mechanical molecular model," *J. Am. Chem. Soc.* **107**, 3902–3909 (1985).
- <sup>83</sup>X. Lopez and D. M. York, "Parameterization of semiempirical methods to treat nucleophilic attacks to biological phosphates: AM1/d parameters for phosphorus," *Theor. Chem. Acc.: Theory, Comput. Modeling* **109**, 149–159 (2003).
- <sup>84</sup>K. Nam, Q. Cui, J. Gao, and D. M. York, "Specific reaction parametrization of the AM1/d Hamiltonian for phosphoryl transfer reactions: H, O, and P atoms," *J. Chem. Theory Comput.* **3**, 486–504 (2007).
- <sup>85</sup>S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu," *J. Chem. Phys.* **132**, 154104–154123 (2010).
- <sup>86</sup>S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory," *J. Comput. Chem.* **32**, 1456–1465 (2011).
- <sup>87</sup>M. Gaus, A. Goez, and M. Elstner, "Parameterization and benchmark of DFTB3 for organic molecules," *J. Chem. Theory Comput.* **9**, 338–354 (2013).
- <sup>88</sup>M. Gaus, X. Lu, M. Elstner, and Q. Cui, "Parameterization of DFTB3/3OB for sulfur and phosphorus for chemical and biological applications," *J. Chem. Theory Comput.* **10**, 1518–1537 (2014).
- <sup>89</sup>E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Banwarth, and S. Grimme, "A generally applicable atomic-charge dependent London dispersion correction," *J. Chem. Phys.* **150**, 154122 (2019).
- <sup>90</sup>M. R. Shirts and J. D. Chodera, "Statistically optimal analysis of samples from multiple equilibrium states," *J. Chem. Phys.* **129**, 124105 (2008).
- <sup>91</sup>M. Ceriotti, D. E. Manolopoulos, and M. Parrinello, "Accelerating the convergence of path integral dynamics with a generalized Langevin equation," *J. Chem. Phys.* **134**, 084104 (2011).
- <sup>92</sup>M. Ceriotti, G. Bussi, and M. Parrinello, "Colored-noise thermostats à la carte," *J. Chem. Theory Comput.* **6**, 1170–1180 (2010).
- <sup>93</sup>M. Ceriotti, G. Bussi, and M. Parrinello, "Langevin equation with colored noise for constant-temperature molecular dynamics simulations," *Phys. Rev. Lett.* **102**, 020601 (2009).
- <sup>94</sup>T. J. Giese, J. Zeng, and D. M. York, "Multireference generalization of the weighted thermodynamic perturbation method," *J. Phys. Chem. A* **126**, 8519–8533 (2022).
- <sup>95</sup>Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang, and E. Weinan, "DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models," *Comput. Phys. Commun.* **253**, 107206 (2020).
- <sup>96</sup>D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 (Accessed Jan 30, 2017).
- <sup>97</sup>K. Nishimoto and N. Mataga, "Electronic structure and spectra of some nitrogen heterocycles," *Z. Phys. Chem.* **12**, 335–338 (1957).
- <sup>98</sup>K. Ohno, "Some remarks on the Pariser–Parr–Pople method," *Theor. Chim. Acta* **2**, 219–227 (1964).
- <sup>99</sup>G. Klopman, "A semiempirical treatment of molecular structures. II. Molecular terms and application to diatomic molecules," *J. Am. Chem. Soc.* **86**, 4550–4557 (1964).
- <sup>100</sup>K. Kosugi, H. Nakano, and H. Sato, "SCC-DFTB-PIMD method to evaluate a multidimensional quantum free-energy surface for a proton-transfer reaction," *J. Chem. Theory Comput.* **15**, 4965–4973 (2019).
- <sup>101</sup>M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, 2010).
- <sup>102</sup>J. R. Cendagorta, H. Shen, Z. Bačić, and M. E. Tuckerman, "Enhanced sampling path integral methods using neural network potential energy surfaces with application to diffusion in hydrogen hydrates," *Adv. Theory Simul.* **4**, 2000258 (2021).
- <sup>103</sup>A. Lamaire, M. Cools-Ceuppens, M. Bocus, T. Verstraelen, and V. Van Speybroeck, "Quantum free energy profiles for molecular proton transfers," *J. Chem. Theory Comput.* **19**, 18–24 (2023).
- <sup>104</sup>T. J. Giese and D. M. York, "Estimation of frequency factors for the calculation of kinetic isotope effects from classical and path integral free energy simulations," *J. Chem. Phys.* **158**, 174105–174115 (2023).
- <sup>105</sup>Z. Zhang, X. Liu, Z. Chen, H. Zheng, K. Yan, and J. Liu, "A unified thermostat scheme for efficient configurational sampling for classical/quantum canonical ensembles via molecular dynamics," *J. Chem. Phys.* **147**, 034109 (2017).
- <sup>106</sup>Z. Zhang, X. Liu, K. Yan, M. E. Tuckerman, and J. Liu, "Unified efficient thermostat scheme for the canonical ensemble with holonomic or isokinetic constraints via molecular dynamics," *J. Phys. Chem. A* **123**, 6056–6079 (2019).

<sup>107</sup>C. W. Hopkins, S. Le Grand, R. C. Walker, and A. E. Roitberg, “Long-time-step molecular dynamics through hydrogen mass repartitioning,” *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).

<sup>108</sup>D. Stanzione, J. West, R. T. Evans, T. Minyard, O. Ghattas, and D. K. Panda, “Frontera: The evolution of leadership computing at the national science foundation,” in *PEARC'20 Practice and Experience in Advanced Research Computing* (Association for Computing Machinery, 2020), pp. 106–111.

<sup>109</sup>J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, and N. Wilkins-Diehr, “XSEDE: Accelerating scientific discovery,” *Comput. Sci. Eng.* **16**, 62–74 (2014).

<sup>110</sup>T. J. Boerner, S. Deems, T. R. Furlani, S. L. Knuth, and J. Towns, “PEARC'23: Practice and experience in advanced research computing,” in *ACCESS: Advancing Innovation: NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services and Supports* (Association for Computing Machinery, 2023), pp. 173–176.