

Software infrastructure for next-generation QM/MM- Δ MLP force fields

Timothy J. Giese, Jinzhe Zeng, Lauren Lerew, Erika McCarthy, Yujun Tao, Şölen

Ekesan, and Darrin M. York*

*Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine
and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ
08854, USA*

E-mail: Darrin.York@rutgers.edu

ABSTRACT: We present software infrastructure for the design and testing of new quantum mechanical/molecular mechanical and machine-learning potential (QM/MM- Δ MLP) force fields for a wide range of applications. The software integrates Amber’s molecular dynamics simulation capabilities with fast, approximate quantum models in the xtb package and machine-learning potential corrections in DeePMD-kit. The xtb package implements recently developed density-functional tight-binding quantum mechanical models with multipolar electrostatics and density-dependent dispersion (GFN2-xTB), and the interface with Amber enables their use in periodic boundary QM/MM simulations with linear-scaling QM/MM particle-mesh Ewald electrostatics. The accuracy of the semiempirical models is enhanced by including machine-learning correction potentials (Δ MLPs) enabled through an interface with the DeePMD-kit software. The goal of this paper is to present and validate the implementation of this software infrastructure in molecular dynamics and free energy

*To whom correspondence should be addressed

simulations. The utility of the new infrastructure is demonstrated in proof-of-concept example applications. The software elements presented here are open-source and freely available. Their interface provides a powerful enabling technology for the design of new QM/MM- Δ MLP models for studying a wide range of problems, including biomolecular reactivity and protein-ligand binding.

Introduction

Molecular simulations provide atomic-level insight to complex chemical phenomena and are indispensable in many areas of fundamental research, technology development, and molecular design.^{1,2} The predictions enabled by molecular simulations are critically dependent on the accuracy of the atomic forces that drive the dynamics.¹⁻⁴ These forces are determined as the negative gradients with respect to the atomic positions of a model potential energy function, commonly referred to as the “force field”. A wide range of force fields exist for performing molecular simulations for different types of applications.⁵⁻⁸ Two particularly challenging areas for force field development are the design of potential energy models that can address problems that involve: (1) chemical reactions in complex heterogeneous condensed phase environments,⁹ and (2) protein-ligand binding free energies used in drug discovery.¹⁰ The former requires an accurate model capable of describing covalent bond cleavage/formation and the associated change in electronic structure that occurs along the reaction’s minimum free energy path. The latter requires quantitatively accurate prediction of intermolecular interactions experienced by candidate drug-like molecules – for which little or no experimental data may exist – in different environments. The challenge in developing accurate force fields for these target areas is further amplified by the large amount of required sampling. This requirement demands that the methods be sufficiently fast to be practical (and hence preclude the use of computationally intensive *ab initio* quantum mechanical methods).¹¹ One promising strategy is to develop generalized hybrid quantum mechanical/molecular

mechanical (QM/MM) models based on fast, approximate semiempirical/density-functional tight-binding QM methods that are augmented by machine-learning potential (MLP) corrections to enhance accuracy while maintaining low computational cost.¹² Combined QM/MM- Δ MLP models have seen a burst in development over the past few years.^{13–17} A critical barrier to progress is the availability of open-source software that enables the integration of the latest QM, MM, and MLP potentials within a molecular dynamics framework that support a wide range of free energy and enhanced sampling methods under periodic boundary conditions with rigorous long-ranged electrostatics. The current work addresses this issue with the integration of software infrastructure to develop new QM/MM- Δ MLP models for use in molecular dynamics and free energy simulations.

The remainder of the manuscript proceeds as follows. The Methods section describes key background and technical aspects of the software integration between Amber,^{18,19} xtb²⁰ and DeePMD-kit,^{21–23} as well as the computational details for the demonstration applications. The Results and Discussion section is organized into two topics. The first topic focuses on applications that demonstrate the effectiveness of the new sander/xtb interface for simulating catalytic reactions using GFN2-xTB QM/MM. Specifically, we apply the new interface to evaluate free energy profiles of the self-alkylation mechanism in artificially engineered RNA enzymes^{25–28} calculated from GFN2-xTB QM/MM umbrella sampling.²⁹ The reaction profiles are compared to DFTB3/3ob and *ab initio* PBE0/6-31G* QM/MM sampling. Furthermore, we demonstrate that inclusion of a Δ MLP correction improves the agreement between the *ab initio* and semiempirical free energy profiles of methyltransferase ribozyme.^{27,28,30} The second topic addressed in the Results and Discussion section focuses on the development of a GFN2-xTB QM/MM- Δ MLP model for drug discovery applications using the new interface between Amber and DeePMD-kit. The ability of the preliminary model to reproduce benchmark relative energies from a variety of databases is demonstrated, and the results are compared to several semiempirical, *ab initio*, and machine learning potentials. Finally, the model is applied to calculate the relative solvation free energies of drug-like tautomers using

an indirect thermodynamic approach.³¹ These calculations demonstrate the ability to correct end-state free energies from molecular mechanical (MM) to quantum mechanical (QM) levels, as would be used in drug discovery applications.

Methods

The framework for the QM/MM- Δ MLP force fields in Amber use the following general form.

$$E_{\text{QM/MM-}\Delta\text{MLP}} = E_{\text{MM}} + E_{\text{QM}} + E_{\text{QM/MM}} + E_{\text{ML}} \quad (1)$$

E_{MM} and E_{QM} are the MM and QM contributions to the energy, and $E_{\text{QM/MM}}$ is their interaction, calculated with electrostatic embedding. E_{ML} is a non-electrostatic correction obtained from a machine learning potential (sometimes referred to as the machine learning potential correction, Δ MLP). In the present work, we extend the capability of the sander program to calculate E_{QM} and $E_{\text{QM/MM}}$ with the GFN2-xTB²⁹ model provided by the xtb package.²⁰ Furthermore, we introduce E_{ML} into sander via an interface to the DeePMD-kit software. The interfaces are described in detail in the next sections.

QM/MM integration of xtb with Amber

The QM/MM interface between Amber and external QM software has often been performed through a file-based process in which a QM input file is written at each MD time step, a system call is made to invoke the QM program, and the energy and forces are read by parsing the QM output file. The xtb software^{20,29,32,33} can be compiled as a library and directly linked to the MD program thereby avoiding the need to make system calls and perform disk operations at each step of dynamics. Furthermore, the direct linkage provides an opportunity to incorporate long-range Ewald electrostatics within the self-consistent field (SCF) procedure.^{24,34–36}

The QM/MM interaction is performed with electrostatic embedding.

$$\begin{aligned}
E_{\text{QM/MM}} = & E_{\text{QM/MM,LJ}} + \sum'_{b \in \text{MM}} \int \int \frac{q_{\text{QM}}(\mathbf{r}') q_b(\mathbf{r} - \mathbf{r}_b)}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \\
& + \sum_{a \in \text{QM}} q_a \left(\Delta\phi_{\text{MM}}(\mathbf{r}_a) + \frac{1}{2} \Delta\phi_{\text{QM}}(\mathbf{r}_a) \right) - \pi \frac{Q_{\text{QM}}^2 + 2Q_{\text{QM}}Q_{\text{MM}}}{2\beta^2 V}
\end{aligned} \tag{2}$$

$E_{\text{QM/MM,LJ}}$ is the nonelectrostatic Lennard-Jones contribution. The second term is the short-range electrostatic interaction between the QM charge density $q_{\text{QM}}(\mathbf{r})$ and the nearby MM charges $q_b(\mathbf{r} - \mathbf{r}_b)$, where the primed summation includes only those MM atoms within a cutoff of the QM region. The remaining terms model the long-range electrostatic interactions. q_a is the Mulliken charge of atom a , $\Delta\phi_{\text{MM}}(\mathbf{r}_a)$ is the electrostatic potential of the MM atoms outside of the cutoff, and $\Delta\phi_{\text{QM}}(\mathbf{r}_a)$ is the long-range interaction between the QM region and its periodic images. The last term is a contribution arising from a neutralizing uniform background correction. The MM energy contains a similar correction such that the net contribution to the total energy depends on the total charge of the system, $-\pi(Q_{\text{MM}} + Q_{\text{QM}})^2/(2\beta^2 V)$.

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

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

The first term in eqs. 3 and 4 is the reciprocal space potential of the MM and QM charge densities, respectively. $\mathbf{k}$ is the angular wave number of the plane wave, V is the volume

of the unit cell, and β is the Ewald coefficient. The second term in eqs. 3 and 4 is a real-space correction that removes the potential of the Ewald Gaussians within a cutoff of the QM region. In many applications the real-space correction appears as a term involving the complimentary error function, which replaces the Ewald Gaussian potential with a point charge potential; however, E_{QM} already contains a model for the QM/QM electrostatic interactions, and the nearby QM/MM interactions are given by the second term in eq. 2.

The software interface is written such that sander, an MD engine within AmberTools,^{18,19} is responsible for calculating E_{MM} , $E_{\text{MM,LJ}}$, $\Delta\phi_{\text{MM}}$, and $\Delta\phi_{\text{QM}}$; and xtb is responsible for calculating E_{QM} , q_a , and the second term in eq. 2. An analogous separation of responsibilities occurs when calculating atomic forces. Before the SCF starts, sander evaluates $\Delta\phi_{\text{MM}}$ with the particle mesh Ewald (PME) method and it precomputes the exponentials appearing in eq. 4. The xtb calculation is initialized from the QM atom positions and the location, charges, and “hardness” values, η , of each MM atom within a cutoff of the QM region. 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.

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

In our interface, one can set a parameter g to control the MM hardness values according to eq. 6, where $\eta(Z)$ is the hardness of the element with atomic number Z .

$$\eta_a = \begin{cases} g, & \text{if } g > 0 \\ \eta(1), & \text{if } g = 0 \\ |g|\eta(Z_a), & \text{if } g < 0 \end{cases} \quad (6)$$

In the present work, we set $g = 0$ which assigns all MM charges the hardness of hydrogen.

We take advantage of the object oriented programming features within the Fortran 2008 language to update and communicate the contribution of the long-range electrostatics within the SCF procedure by creating a new class that inherits from the `TSolvation` data structure defined within `xtb`. This structure is normally used to variationally interact the QM region with an implicit solvent model; however, the resulting child class redefines the class methods to evaluate the last term in eq. 2 and the potential shift of each QM atom: $\Delta\phi_{\text{MM}}(\mathbf{r}_a) + \Delta\phi_{\text{QM}}(\mathbf{r}_a)$. We have validated the `sander`/`xtb` interface and the incorporation of Ewald electrostatics into the `xtb` self-consistent field procedure by verifying that the total energy is conserved in simulations performed in the microcanonical ensemble. Details regarding this test can be found in the Supporting Information.

QM/MM integration of DeePMD-kit with Amber

We have interfaced the DeePMD-kit library^{21–23} to the `sander` MD package.¹⁹ The implementation was heavily influenced by the development of so-called Δ MLP potentials, which use neural networks to correct an underlying potential energy function^{12,37–41} as in eq. 1. Specifically, we are interested in using new range-corrected semiempirical QM/MM- Δ MLP potentials that correct the semiempirical QM/MM energies and forces to better reproduce *ab initio* QM/MM.⁴¹ From a user’s perspective, one provides a neural network parameter file and selects a set of atoms whose energies and forces should be corrected. This selection normally corresponds to the QM atoms in the system, such that the semiempirical QM/QM interactions are corrected with the Δ MLP. Furthermore, one can optionally set a cutoff radius that will calculate a Δ MLP correction between the QM atoms and nearby MM atoms within the cutoff.⁴¹ The DPRc neural network parameter files include settings that smoothly turn off the Δ MLP QM/MM corrections over a range (typically from 5-to-6 Å), and the cutoff radius should either be 0 (no QM/MM corrections) or the upper-bound of the range (typically 6 Å). When `sander` requests the energy and force corrections from DeePMD-kit, it provides the position, atom type, and residue number of each QM atom and those MM

atoms within a cutoff of the QM region. From the perspective of DeePMD-kit, the QM atom types are their 2 character element symbols. The MM atom types are the letter “m” followed by the 2 character element symbol. This strategy allows the neural network to correct a QM/QM interaction differently than a QM/MM interaction.

The DPRc energy is a sum of atomic energy contributions, E_i :

$$E_{\text{DPRc}} = \sum_{i=1}^N E_i(\mathbf{r}_i, \{\mathbf{r}_j\}_{j \in n(i)}) \quad (7)$$

where N is the number of atoms, $\mathbf{r}_i$ is the position of atom i , and $n(i)$ is the set of neighboring atom indexes within a cutoff of atom i (the DPRc cutoff). The atomic decomposition of the DPRc potential (eq. 7) model makes it well-suited for parallel computation. Expressions for the neural networks used to calculate the E_i values can be found elsewhere.^{12,41}

QD π -2 QM/QM model

We describe the preliminary training of a QD π -2 force field based on GFN2-xTB²⁹ as a motivating demonstration of the new QM/MM- Δ MLP software infrastructure for drug discovery applications. The model is preliminary in that only the QM/QM interactions have been trained; it remains to tune the QM/MM interactions with a separate neural network model.

The QD π -2 QM/QM model is a Δ MLP correction to the GFN2-xTB²⁹ energies and forces.

$$E_{\text{QD}\pi\text{-2 QM/QM}} = E_{\text{GFN2-xTB}} + E_{\text{DPRc}} \quad (8)$$

The DPRc model supplements the deep-potential model used in the original QD π ⁴² by introducing the attention-based descriptor⁴³ with 2 attention layers, as described in Ref. 23.

To train the model, we prepared target data from several databases, including: SPICE,⁴⁴ ANI-1x⁴⁵ and ANI-2x⁴⁶ (ANI), GEOM drug,⁴⁷ FreeSolv,⁴⁸ and the relative energy (RE) datasets collected in the previous QD π model.⁴² The entirety of the SPICE and RE datasets

were used in the initial training to produce 4 sets of neural network parameters, because we had already recalculated the reference values at a consistent ω B97M-D3BJ/def2-TZVPPD level of theory. The ANI-1x, ANI-2x, and GEOM datasets were not tabulated at the same level of theory, and recalculating the entire databases would be time-consuming; therefore, we used an active learning strategy implemented in the DP-GEN program⁴⁹ to reduce the number of ω B97M-D3BJ/def2-TZVPPD evaluations with these other datasets. The energy and force corrections for each molecule in the ANI-1x, ANI-2x, and GEOM datasets were calculated from the 4 initial parameter sets. If the standard deviation in the energies and forces exceed a threshold, then a reference ω B97M-D3BJ/def2-TZVPPD calculation of the molecule was performed and included in the training data. Specifically, a structure is selected for labeling when the standard deviation in the energy-per-atom exceeds $\sigma_E \geq 0.015$ eV/atom or the maximum standard deviation in the atomic force exceeds $\sigma_F \geq 0.20$ eV/Å. Four new sets of neural network parameters were optimized with the expanded set of target data, the databases are recalculated with each model, and additional data is selected for labeling. The active learning procedure terminates when no new target data is selected for labeling. The target data was similarly expanded by using an active learning strategy⁴¹ to collect data from 10 ps GFN2-xTB QM/MM simulations of each molecule in the FreeSolv database (FreeSolv-MD). The simulations were performed in explicit OPC water,⁵⁰ and the solute configuration was saved every 1 ps. The corrections produced by the neural network parameters sets were compared, and the structure was labeled with ω B97M-D3BJ/def2-TZVPPD if the deviation between the 4 models was large. The target data was further expanded from 1 ps gas phase simulations of the RE monomers (RE-MD). Upon termination of the active learning procedure, the expanded set of target data is partitioned into training and validation subsets. The neural network parameters are re-optimized using the training data, which consists of 95% of the expanded target data. By following this procedure, we collected a total of 997,535 (SPICE), 331,341 (ANI), 23,569 (GEOM), 31,163 (FreeSolve-MD), 12,019 (RE-MD), 175,193 (RE) data points. The data was divided into training and validation subsets, where 5% of

the data was reserved for validation. The entirety of the RE dataset was included in the training because samples from RE-MD was included in the validation subset.

We used the PSI4 program⁵¹ to perform the ω B97M-D3BJ/def2-TZVPPD labeling.⁵² Comparisons will be made against the COMP6 database.⁵³ This database was not included in the training nor validation sets. We recalculated the COMP6 database at the ω B97M-D3BJ/def2-TZVPPD level of theory to make comparison. The QD π -2 neural network parameters were trained for 18 million steps using the Adam optimizer with a batch size of 16.

The differences between the QD π -1 and QD π -2 models are summarized:⁴²

1. The number of elements is expanded from 4 (H,C,N,O) to 12 (H,C,N,O,F,Na,P,S,Cl,K,Br,I).
2. The target QM level is upgraded from ω B97X/6-31G*⁵⁴ to ω B97M-D3BJ/def2-TZVPPD.⁵²
3. The base semiempirical QM model is changed from DFTB3/3ob^{55,56} to GFN2-xTB.²⁹
4. The Δ MLP model is enhanced using the attention-based descriptor.⁴³
5. The training data is expanded by using an active learning procedure, as previously described.

The datasets are used to make comparison between several potentials, including: ω B97X/6-31G*⁵⁴ using PSI4,⁵¹ GFN2-xTB²⁹ using xtb,²⁰ DFTB3/3ob^{55,56} using DFTB+,⁵⁷ a pure Deep Potential model trained against the SPICE dataset for 10M steps (DP-SPICE),²³ and other methods that were benchmarked in our previous work.⁵⁸

QM/MM- Δ MLP alchemical free energy calculations

Alchemical free energy (AFE) simulations are typically applied to classical pairwise additive molecular mechanical (MM) force fields.¹⁰ The Amber software has undergone significant improvements to support AFE simulations^{59,60} by introducing efficient optimized pathways,⁶¹ new smoothstep softcore potentials,^{61,62} and λ -scheduling features⁶³ that enable the direct

transformation of one MM molecule into another. These “direct AFE” simulations often require specialized enhanced sampling methods⁶⁴ to reliably converge the free energy estimates. Hence, AFE simulations are more practical when used with computationally efficient MM force fields. Direct AFE simulations using quantum mechanical models, such as QM/MM- Δ MLPs, present much greater challenges because they are not pairwise decomposable and relatively expensive to evaluate.

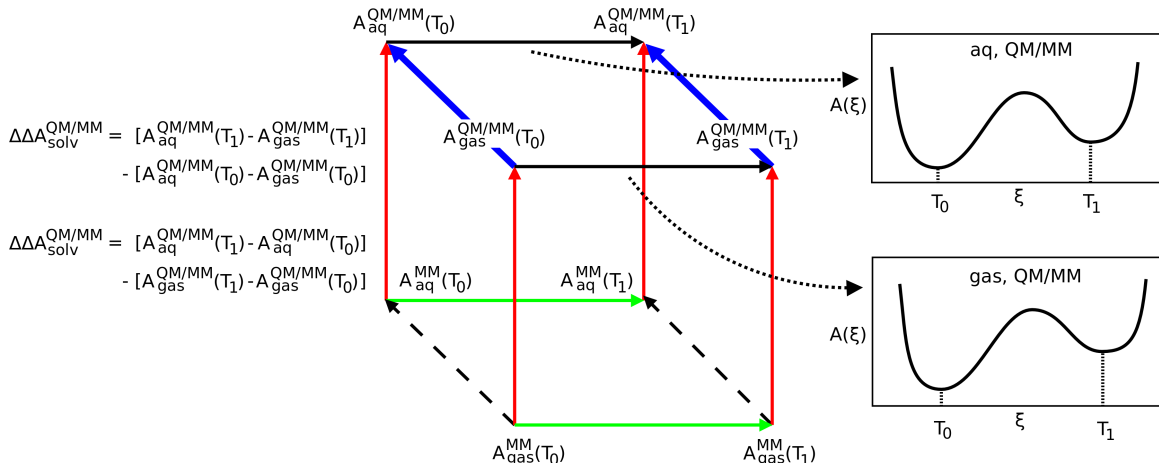


Figure 1: Schematic of two approaches for calculating the relative solvation free energy of tautomers T_0 and T_1 with a QM/MM potential. The two equations (left) are equivalent expressions for the solvation free energy. The top expression is the difference between the QM/MM solvation free energies for each tautomer indicated by the blue arrows shown in the thermodynamic cycle. The bottom expression is the difference between the two black arrows. The thermodynamic cycle depicts the “indirect approach”, whereby the relative solvation free energy at the QM/MM level is evaluated from alchemical transformations using a MM reference potential (the green arrows) and MM→QM/MM “book-end” corrections (the red arrows). Alternatively, the relative solvation free energy can be estimated from a free energy surface (depicted on the right), $A(\xi)$ that physically connects the T_0 and T_1 QM/MM states (the black arrows) using a proton transfer reaction coordinate, ξ .

An alternative is to use an “indirect” approach that utilizes MM force fields to represent intermediate alchemical states, as illustrated in Figure 1. Such calculations require additional end-state calculations to account for the free energy of changing the MM representation of a molecule into a QM/MM- Δ MLP representation (or vice versa). The end-state calculations within the indirect AFE approach are sometimes referred to as “book-ending corrections”.³¹

The indirect AFE approach has the advantage that the computationally intensive alchemical steps can be conducted efficiently using MM force fields and existing free energy infrastructure, whereas the MM→QM book-ending corrections – which do not involve any alchemical changes – can be efficiently performed with only modest sampling requirements.

There have been two widely used methods for performing MM→QM book-ending corrections:¹⁰ free energy perturbation (FEP) analysis of equilibrium sampling^{31,65–68} and non-equilibrium work methods.^{67,69–72} In principle, equilibrium FEP methods can be performed in a single step that requires only MM sampling. In practice, this is usually not possible unless a specifically tuned “reference potential” is used that has sufficient phase space overlap with the high-level QM method. Reliable free energy estimates often require sampling of the high-level QM method and use of Bennett’s acceptance ratio (BAR) method.⁷³ One strategy that has been demonstrated to be successful is to perform short high-level QM simulations. These short simulations are used to parameterize a MM reference potential through force-matching, and the reference potential is extensively sampled. The free energy is then estimated from BAR analysis of the high-level and reference potential sampling.³¹

Non-equilibrium work^{74–76} methods, on the other hand, require equilibrium sampling at either one or both end states. A series of short “switching” simulations, departing from the end-state samples, are used to form a collection of work values. A switching simulation dynamically varies λ from 0 to 1, where $\lambda = 0$ and $\lambda = 1$ represent the MM and QM/MM potentials, respectively, and the free energy difference between the two states is calculated from the collection of work values.

We demonstrate that the Amber software infrastructure is capable of performing both equilibrium FEP and non-equilibrium work book-ending corrections using new QM/MM- Δ MLP force field models that integrate GFN2-xTB and deep-learning potentials from DeePMD-kit. We apply the methods to calculate relative solvation free energies of drug-like tautomers. The 9 tautomer reactions examined in this work are illustrated in Figure 2.

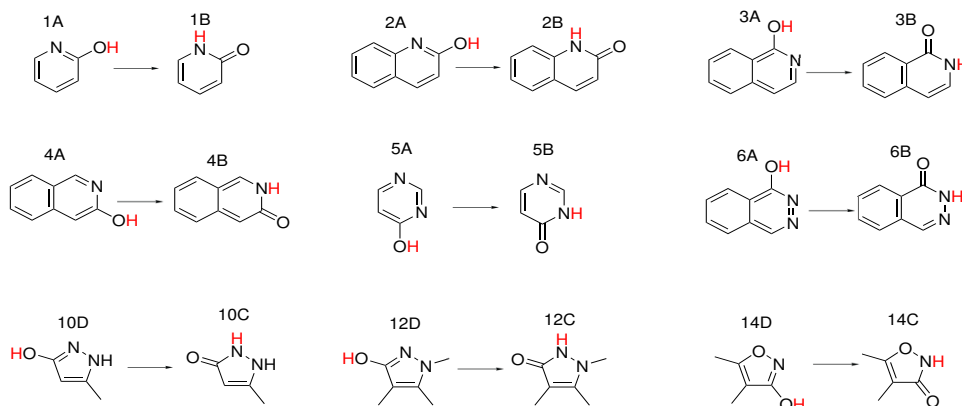


Figure 2: Small molecule tautomeric reactions examined in this work. The red hydrogens highlight the transferred proton.

Computational Details

Free energy surface simulations of artificially engineered RNA enzymes. The new integrated sander/xtb/DeePMD-kit software is applied to simulations of self-alkylation reactions in artificially engineered RNA enzymes, in particular the self-alkylating ribozyme²⁶ (SAR) and the methyltransferase ribozyme^{27,28,30} (MTR1). The SAR and MTR1 systems were built from crystallographic structures PDB ID 6XJW²⁶ and 7V9E,²⁸ respectively. The crystal structures correspond to the product state, and the active site coordinates were manipulated to represent the reactant state. Furthermore, we protonated the N3 position of residue C10 in the MTR1 system. The MTR1 system was solvated with 18250 waters in a truncated octahedron with unit cell lengths of 90.2 Å. A total of 113 Na⁺ and 47 Cl⁻ ions were added to neutralize the system and yield a 140 mM bulk ion concentration. The SAR system was similarly solvated with 15181 waters 96 Na⁺ and 38 Cl⁻ ions in a orthorhombic cell whose average lattice vector length is 79.3 Å.

All simulations were performed with ff99OL3 RNA force field,⁷⁷ TIP4P/Ew water model,⁶ and the monovalent ion parameters developed by Joung and Cheatham.⁷⁸ The MM simulations used the particle mesh Ewald method with 10 Å real-space cutoffs, tinfoil boundary conditions,^{79,80} and a 1 Å reciprocal space grid. The systems were equilibrated in a multistep

procedure described in Ref 81 consisting of 150 ns of dynamics.

The free energy profiles were analyzed from QM/MM umbrella sampling⁸² departing from the MM reactant state. The QM/MM models considered are: DFTB3/3ob, GFN2-xTB, and PBE0/6-31G*. The SAR “baseline” reaction involves nucleophilic attack of the guanine N7 to an epoxide ligand. This reaction was modeled with a neutral QM region consisting of 30 atoms, including the 16 nucleobase and a subset of the epoxide ligand. We also examine an “activated” mechanism whose reactant state includes a protonated epoxide oxygen. In this case, the QM region consisted of 31 atoms with a net 1+ charge. The MTR1 QM region has a 1+ net charge and includes 48 atoms from the O6mG ligand and the C10 and A63 nucleobases. The semiempirical electrostatic calculations were evaluated with the QM/MM Ewald approach introduced by Nam, Gao and York,²⁴ whereas the *ab initio* QM/MM simulations used the ambient potential composite Ewald method.³⁶ In all cases, the Ewald calculations employed 10 Å real-space cutoffs, a 1 Å reciprocal space grid, and tinfoil boundary conditions.^{79,80} The unbiased free energy surface was obtained from solution of the unbinned weighted histogram analysis method (UWHAM)^{83,84} and multistate Bennett acceptance ratio (MBAR) equations⁸⁵ using the FE-ToolKit software.⁸⁶

The SAR mechanism was described with a single reaction coordinate, $\xi_1 = R_{C22-O21} - R_{C22-G16:N7}$, which was sampled from -1.5 Å to 1.0 Å using 32 umbrella windows. The DFTB3/3ob and GFN2-xTB simulations were sampled for 50 ps/window using a 1 fs time step, whereas the PBE0/6-31G* simulations were sampled for 10 ps/window with a 1 fs time step.

The MTR1 free energy surface was generated from GFN2-xTB QM/MM umbrella sampling involving 2 reaction coordinates: $\xi_1 = R_{C10:N3-H} - R_{O6mG:N1-H}$ and $\xi_2 = R_{O6mG:O6-C} - R_{A63:N1-C}$. The minimum free energy path was determined from 50 iterations of the surface-accelerated string method.⁸⁷ Each string was discretized with 32 windows and each window was sampled for 4 ps in each iteration (6.4 ns of aggregate sampling). We compare the GFN2-xTB free energy profile to the DFTB3/3ob and PBE0/6-31G* profiles previously re-

ported in Ref. 88. The umbrella sampling was performed in the canonical ensemble using the Langevin thermostat with a collision frequency of 5 ps⁻¹ to maintain a temperature of 298 K.

Alchemical free energy calculations of drug-like tautomers. We calculate the free energy of 9 tautomeric reactions (Figure 2) taken from tautomeric database.⁸⁹⁻⁹¹ The relative solvation free energy is evaluated from an indirect approach that performs the AFE simulations with a MM model. The end-state free energies are then supplemented with MM-to-QM/MM book-end corrections. We compare two strategies for calculating the book-end corrections. The first strategy analyzes equilibrium sampling produced by a λ -dependent potential energy function, $U(\lambda) = (1 - \lambda)U_{\text{MM}} + \lambda U_{\text{QM/MM}}$, at 21 uniformly spaced λ values. The second strategy estimates the book-end correction from nonequilibrium statistics using Jarzynski’s equation. Furthermore, we compare the relative solvation free energies to results obtained from QM/MM umbrella sampling, which does not involve a MM reference potential. The umbrella sampling is analyzed to produce a free energy surface of the proton transfer. The reaction free energy is the difference between the product and reactant state minima on the surface. The relative solvation free energy is obtained by repeating the calculations in the gas phase and taking the difference from the aqueous phase result. The calculation details are described below. The MM-to-QM/MM bookend simulations and the QM/MM umbrella sampling were performed with the DFTB3/3ob, GFN2-xTB, and QD π -2 models. The QD π -2 model differs from the GFN2-xTB QM/MM calculations only by the inclusion of neural network correction to the QM/QM interactions.

Each molecule shown in Figure 2 was modeled with the GAFF2 force field⁹² and solvated in a truncated octahedron (whose real space lattice vector lengths were approximately 48.5 Å) with 3000 TIP4P/Ew waters.⁶ The solvated systems were equilibrated for 200 ps in the *NPT* ensemble at 1 atm and 298.15 K. The Berendsen barostat and Langevin thermostat were utilized to control the pressure and temperature with a collision frequency of 5 ps⁻¹. The systems were further equilibrated for an additional 100 ps in the canonical ensemble

at 298.15 K. The initial gas phase structures were prepared by removing the solvent and equilibrating the molecule at 298.15 K for 50 ps.

The alchemical free energy calculations were performed with the second-order smoothstep function,⁶² and the alchemical space was uniformly discretized by 21 λ values. The initial configurations were geometry optimized for 5000 steps of steepest descent minimization. This was followed by 2 ps of equilibration and 6 ns of simulation in the canonical ensemble at 298.15 K with a 1 fs time step. Hamiltonian replica exchange attempts were made every 20 fs, and the first 200 ps of simulation was discarded from the analysis. The gas phase AFE simulations were performed analogously, and the free energies were calculated from MBAR analysis.⁸⁵

Equilibrium book-ending corrections were applied to the $\lambda = 0$ and $\lambda = 1$ end-states in both the solvated and gas phase environments. Each MM-to-QM/MM correction used 21 uniformly spaced windows, and each window was simulated for 100 ps at 298.15 K using a 1 fs time step. The first 25% of each simulation was discarded as equilibration, and the free energy was estimated from MBAR analysis.

The book-ending corrections were also calculated from Jarzynski’s equation.^{74,75}

$$\Delta A_{\text{MM} \rightarrow \text{QM}} = -k_B T \ln \left\langle \exp \left(\frac{-W_{\text{MM} \rightarrow \text{QM}}}{k_B T} \right) \right\rangle_{\text{MM}} \quad (9)$$

k_B is Boltzmann’s constant, T is the temperature, and W is the accumulated non-equilibrium work required to change the potential from MM to QM/MM in each switch.

$$W = \int_0^{t_s} \frac{\partial U}{\partial \lambda} \frac{\partial \lambda}{\partial t} dt \quad (10)$$

t_s is the switch time, and the switching rate is constant $\dot{\lambda} = t_s^{-1}$. The exponential in eq. 9 is averaged over starting configurations taken from the MM ensemble. A total of 1600 switching simulations were used to calculate the average, and each switch continuously varied λ over the course of 2 ps. The 1600 configurations were obtained from 16 ns of MM sampling.

The tautomer free energies were also calculated from profiles generated from umbrella sampling of the proton transfer reaction coordinate, $\xi_{\text{PT}} = R_{\text{O-H}} - R_{\text{N-H}}$. The reaction coordinate was scanned with 51 harmonic biasing potentials using $300 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ force constants. Initial configurations were created by scanning the reaction coordinate with a sequence of brief 500 fs simulations, and each window was equilibrated for 10 ps. Production sampling was made in the canonical ensemble at 298.15 K for 100 ps, and each simulation was repeated 4 times using different thermostat random number seeds. The free energy surfaces were calculated from MBAR analysis using the FE-ToolKit software⁸⁶ upon discarding 25% of the data. Error estimates were made from circular moving block bootstrap analysis of the data collected from the 4 trials.

Results and Discussion

Application of the new sander/xtb interface to simulation of catalytic reactions

The study of synthetically engineered nucleic acid enzymes shows promise to significantly improve design, creation, and implementation of superior artificial nucleic acid enzymes.^{98–101} Two important exemplars are the self-alkylating ribozyme²⁶ (SAR) and the methyltransferase ribozyme³⁰ (MTR1). These artificially evolved ribozymes catalyze C-N bond formation, a transformation that has not yet been observed in naturally occurring ribozymes, but would be essential for nucleic acid synthesis and early metabolic transformations within the context of the RNA world hypothesis.¹⁰² We apply the software infrastructure developed in this work to compare *ab initio*, DFTB3/3ob and GFN2-xTB QM/MM free energy profiles of SAR and MTR1.

1D free energy profile for a self-alkylating ribozyme. The SAR (Figure 3a) catalyzes the ring opening of an epoxide ligand, forming an irreversible nitrogen-carbon bond between the ligand and the N7 of a specific guanine (G16), as illustrated in Figure 3b.²⁶ We performed

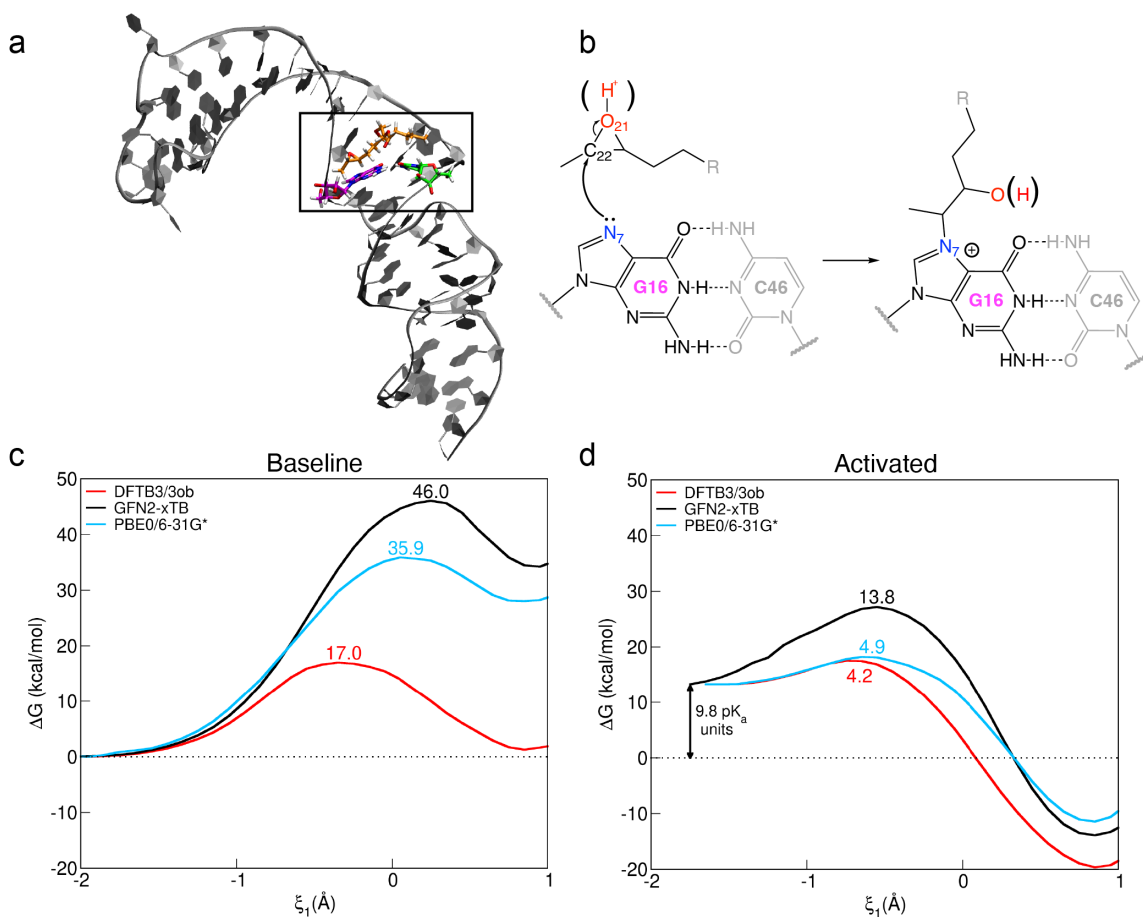


Figure 3: Reaction catalyzed by the self-alkylating ribozyme. (a) The SAR structure. (b) The reaction mechanism involves nucleophilic attack of the guanine N7 to an epoxide ligand. The schematic illustrates the “baseline” mechanism, where the epoxide oxygen is neutral (unprotonated). The “activated” mechanism differs only by having protonated the epoxide oxygen. (c) QM/MM free energy profiles of the “baseline” mechanism. (d) QM/MM free energy profiles of the “activated” mechanism.

QM/MM umbrella sampling to calculate 1-dimensional free energy profiles of this process using GFN2-xTB, DFTB3/3ob, and PBE0/6-31G*. A “baseline” mechanism (Figure 3c) was explored where the nucleophilic attack of N7 occurs on a neutral epoxide, leading to a primary alkoxide product. The “activated” mechanism (Figure 3d) starts with a protonated epoxide to facilitate ring opening and leads to a more stable primary alcohol product. The activated reactant state free energies have been shifted by 13.3 kcal/mol to account for the protonation of the epoxide (-2.8 pK_a) at $pH\ 7$.¹⁰³

All QM/MM models predict the baseline mechanism to have a higher barrier than the activated ligand, as expected. However, the transition state barriers are quantitatively different. DFTB3/3ob produces a baseline mechanism that is nearly isoergic with a 16.7 kcal/mol forward barrier, whereas PBE0/6-31G* and GFN2-xTB predict the product state to be considerably higher in free energy, and their forward barriers are larger (35.9 and 46.0 kcal/mol, respectively). For the activated mechanism, the DFTB3/3ob and PBE0/6-31G* barriers (3.8 and 2.6 kcal/mol, respectively) are similar, whereas GFN2-xTB is markedly higher (13.8 kcal/mol). These results demonstrate that for this reaction DFTB3/3ob and GFN2-xTB give quantitatively different results from one another and from PBE0/6-31G*. This strongly suggests that there is need to further augment these models with ΔMLP corrections to achieve improved accuracy.

2D free energy profile for a methyltransferase ribozyme. The MTR1 catalyzes the nucleophilic attack of adenine (A63) at position N1 to the O⁶-methylguanine (O6mG) ligand at position C6. The attack is assisted by a proton transfer from N3 of C10 to the N1 of O6mG, as illustrated in Figure 4a. We applied the new sander/xtb interface to calculate a 2-dimensional free energy surface of the MTR1 reaction from GFN2-xTB QM/MM umbrella sampling (Figure 4c). The GFN2-xTB minimum free energy path is predicted to be step-wise, in agreement with previous work using DFTB3/3ob and PBE0/6-31G*.⁸⁸ A proton is transferred from C10:N3 to O6mG:N1, and this is followed by a methyl transfer from O6mG:O6 to A63:N1. The inset of Figure 4c compares the GFN2-xTB, DFTB3/3ob, and

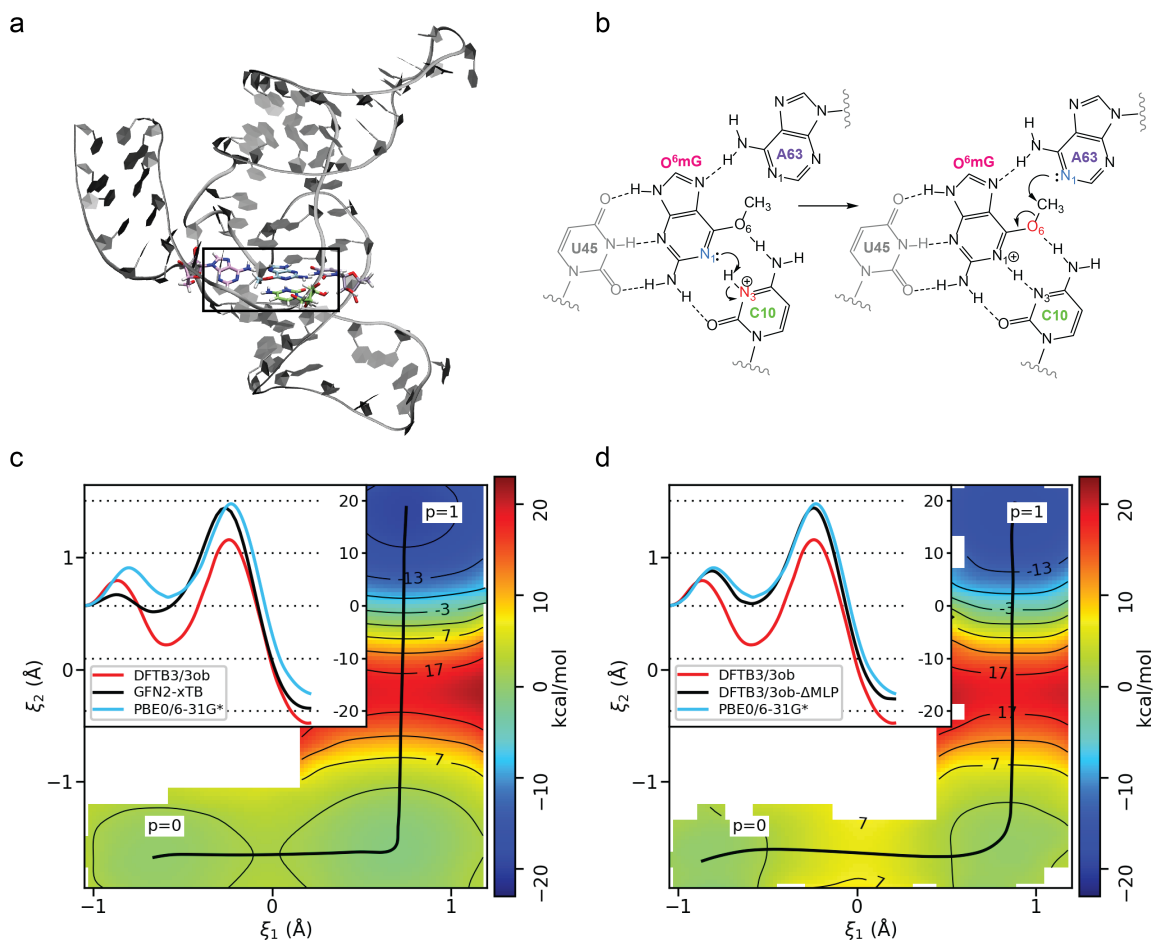


Figure 4: Reaction catalyzed by the methyltransferase ribozyme. (a) The MTR1 structure. (b) The MTR1 mechanism involves nucleophilic attack of adenine (A63) N1 at the O⁶-methylguanine (O6mG) ligand C6 position. The attack is assisted by a proton transfer from N3 of C10 to N1 of O6mG. (c) The 2-dimensional GFN2-xTB QM/MM free energy surface and minimum free energy path. The inset compares the minimum free energy profile with DFTB3/3ob and PBE0/6-31G* QM/MM umbrella sampling results taken from Ref. 88. (d) The 2-dimensional DFTB3/3ob- Δ MLP QM/MM free energy surface and minimum free energy path. The $p = 0$ and $p = 1$ labels denote the reactant and product states, respectively.

PBE0/6-31G* free energy profiles. The profiles are qualitatively similar, but quantitatively different. The activation free energy predicted by GFN2-xTB (18.6 kcal/mol) agrees much better with PBE0/6-31G* (19.4 kcal/mol) than DFTB3 (15.5 kcal/mol). Nonetheless, the barrier for the proton transfer step using GFN2-xTB (2.1 kcal/mol) is considerably lower than PBE0/6-31G* (7.2 kcal/mol). The DFTB3 profile exhibits an overstabilization of the protonated C10 intermediate (-7.5 kcal/mol) relative to *ab initio* (1.6 kcal/mol). The extent of overstabilization is not as severe with GFN2-xTB (-1.1 kcal/mol), but it is still significant. The experimental barrier of the methyl transfer reaction is approximately 21.1 kcal/mol, based on the intrinsic rate of reaction.²⁷ All of the methods explored here underestimate the barrier. The PBE0/6-31G* and GFN2-xTB barriers are within 1.7 and 2.5 kcal/mol of this estimate, respectively, whereas the DFTB3 prediction is lower by 5.6 kcal/mol.

The computational cost of PBE0/6-31G* QM/MM simulations quickly becomes prohibitive to apply. Although the DFTB3 and GFN2-xTB QM models provide a practical alternative in terms of computational efficiency, we find there are important quantitative differences from the PBE0/6-31G* results. For this system, the performance of the GFN2-xTB method is competitive with DFTB3; either method may serve as a useful base model to construct new QM/MM- Δ MLP potentials. As a proof of concept demonstration, Figure 4d illustrates the MTR1 free energy surface calculated from umbrella sampling performed with a preliminary DFTB3/3ob- Δ MLP model trained to reproduce the PBE0/6-31G* QM/MM energies and forces. The Δ MLP uses the DPRc potential,⁴¹ and it is trained in a manner similar to our previous works.^{12,41,104} In brief, the DPRc model consists of filtering and fitting neural networks. The fitting network consists of 3 hidden layers (240 neurons/layer) that transform an atomic descriptor matrix to an atomic energy contribution to the MLP. In the current work, we use the two-body embedding smooth edition deep potential descriptor,²³ which is calculated from the filtering network consisting of 3 hidden layers (25, 50, and 100 neurons). The input to the filtering network is a “coordinate matrix” that describes an atom’s environment. This description includes switching functions that smoothly eliminate correc-

tions between QM and MM atoms as a function of distance such that there is no correction beyond 6 Å. The training consisted of 10 iterations of active learning with 4 MLP parameter sets. The parameters were optimized from 100,000 steps of the Adam algorithm¹⁰⁵ using an exponentially decaying learning rate that varied from 10^{-3} to 5×10^{-8} . Each round of active learning performed umbrella sampling with 224 umbrella windows covering the relevant areas of the 2-dimensional free energy surface. The simulations were sampled for 1 ps with a 1 fs time step, and the active learning procedure incorporated approximately 10,000 PBE0/6-31G* samples into the training procedure, in total. After training, the minimum free energy path of the DFTB3/3ob- Δ MLP was found using the finite temperature string method,⁸⁷ and production simulations were performed using the same protocol described for the GFN2-xTB umbrella sampling.

The DFTB3/3ob- Δ MLP model shows improved agreement with PBE0/6-31G* than the uncorrected GFN2-xTB and DFTB3/3ob semiempirical models. The proton transfer barrier (6.9 kcal/mol) and rate-limiting transition state barrier (19.1 kcal/mol) are within 0.4 kcal/mol of the PBE0/6-31G* values. In this particular application, the Δ MLP was trained for a specific reaction. Although the scope of applicability is limited, it provides access to simulation time scales that would not be feasible with *ab initio* QM/MM sampling. Similarly, it makes it possible to perform path integral molecular dynamics, which requires multiple force evaluations at each time step. In the next section, we highlight the use of Δ MLPs for modeling a wider chemical space for drug discovery applications.

Proof of concept demonstration in drug discovery applications

In this section, we show that a GFN2-xTB- Δ MLP model can be trained to achieve quantitative accuracy with respect to a reference high-level *ab initio* method. Specifically, we focus the demonstration on the creation of a preliminary model for improving QM/QM interactions. The parametrization of a Δ MLP to improve QM/MM interactions will require further training. The QM/MM- Δ MLP presented here is a second-generation version of the

Quantum Deep-learning Potential Interaction model,^{42,58} designated QD π -2.

Preliminary GFN2-xTB/ Δ MLP model for drug-like molecules

We compare energy and force mean absolute errors (MAEs) relative to ω B97M-D3BJ/def2-TZVPPD against a broad set of databases using GFN2-xTB/ Δ MLP and several other models. Table 1 summarizes the MAEs produced by QD π -2 on the training data. Furthermore, the MAEs of QD π -2, ω B97X/6-31G*, GFN2-xTB, DFTB3/3ob ("DFTB3"), and DP-SPICE methods are compared using the validation data. The QD π -2 model produces similar errors between the training and validation data, suggesting the model is not overfit. QD π -2 outperforms the other models on all validation datasets; the GFN2-xTB/ Δ MLP model is more accurate than GFN2-xTB, the pure MLP model, and even a low-level DFT method. The results suggest that GFN2-xTB is a more accurate baseline QM model than DFTB3 for these datasets. This observation is consistent with previous findings.^{42,58} Table 2 compares force and relative energy MAEs using subsets of the ANI, COMP6, and RE databases. The subsets are limited to molecules containing the elements: C, H, O, and N. Similar comparisons were reported in Ref 58 using ω B97X/6-31G* reference data; however, we recalculated the reference data in the present work with a higher level of theory (ω B97M-D3BJ/def2-TZVPPD). Consequently, the models previously trained to reproduce ω B97X/6-31G* calculations do not appear as accurate in Table 2. The QD π -2 model is the most accurate method for all datasets. The GFN1-xTB and GFN2-xTB models are the most accurate semiempirical models.

The accurate prediction of tautomer and protonation states is important for drug discovery applications because 30% of the compounds in vendor databases and 21% of drug databases have potential tautomers.^{106,107} Furthermore, it has been estimated that up to 95% of drug molecules contain ionizable groups.¹⁰⁶ Figures 5 and 6 compare relative protonation energies and tautomerization energies, respectively. These comparisons include QD π -2, ω B97X/6-31G*, GFN2-xTB, and DFTB3. The reference values were calculated at

the ω B97M-D3BJ/def2-TZVPPD level of theory. The QD π -2 model is the most accurate, followed by ω B97X/6-31G*, GFN2-xTB, and DFTB3. This illustrates that machine-learning corrections can be used to achieve quantitative accuracy at an affordable computational cost.

Table 1: Mean absolute error (MAE) of the energies (in kcal/mol) and forces (in kcal/mol/Å) from the training and validation data.^a

Dataset	Training		Validation						
	QD π -2		QD π -2		ω B97X	GFN2-xTB	DFTB3	DP-SPICE	
	E	F	E	F	F	F	F	E	F
SPICE	1.19	1.36	1.19	1.36	2.40	3.98	6.02	7.75	2.62
ANI	1.51	2.28	1.55	2.29	3.15	6.75	11.09	14.22	11.44
GEOM	2.22	1.42	1.80	1.32	2.61	3.53	7.03	16.78	3.05
FreeSolv-MD	0.72	1.39	0.71	1.41	3.17	5.18	8.25	27.14	6.85
RE-MD	1.55	1.69	1.26	1.58	2.50	3.92	8.60	5.77	3.15

^a The datasets include SPICE,⁴⁴ ANI,^{45,46} GEOM,⁴⁷ FreeSolv-MD,⁴⁸ and RE-MD.⁴¹ The validation data compares ω B97M-D3BJ/def2-TZVPPD⁵² using PSI4⁵¹ and QD π -2 using xtb²⁰ and DeePMD-kit,²³ ω B97X/6-31G*¹⁰⁸ (ω B97X) using PSI4,¹⁰⁹ GFN2-xTB²⁹ using xtb,²⁰ DFTB3/3ob^{55,56} using DFTB+,⁵⁷ and the pure Deep-Potential model trained against the SPICE dataset²³ (DP-SPICE).

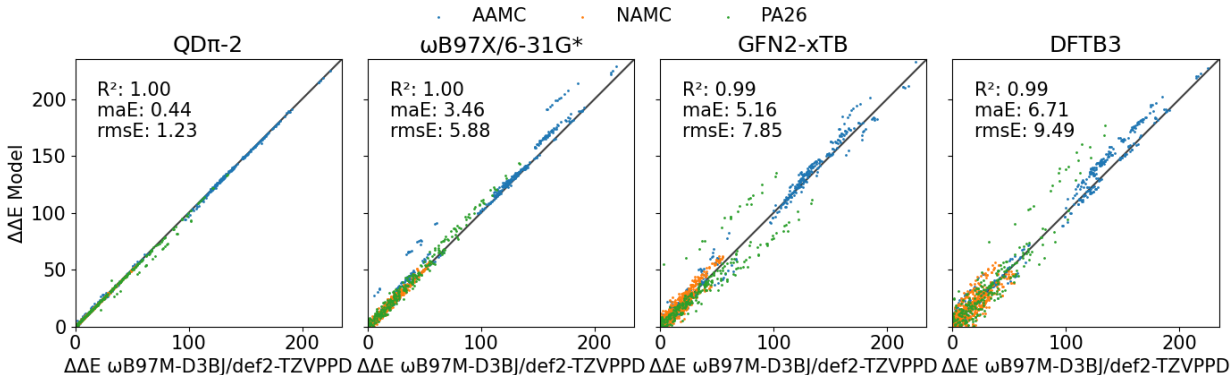


Figure 5: Comparison of relative protonation energies (kcal/mol). The reference values are ω B97M-D3BJ/def2-TZVPPD. The comparisons include molecules from the AAMC,¹¹⁶ NAMC,¹¹⁶ and PA26¹¹⁷ databases.

Table 2: Relative energy (in kcal/mol) and force (kcal/mol/Å) mean absolute errors calculated from subsets of the ANI (ANI-1x and ANI-2x), COMP6 (S66 and COMP5), and RE (TB, AA, NA, PA, HB, T15, SQM) databases.^a

Method	ANI F	S66 ΔE	TB ΔE	AA ΔE	NA ΔE	PA ΔE	COMP5 F	HB ΔE	T15 ΔE	SQM ΔE
QD π -1	4.51	1.14	2.82	6.61	1.94	2.38	3.11	0.74	2.53	3.16
AIQM1	3.45	1.58	1.29	4.60	2.22	5.04	2.08	0.68	1.04	2.84
ANI-1x	6.32	1.94	3.51	84.29	9.64	35.40	4.29	1.33	3.56	16.57
ANI-2x	4.50	1.30	3.08	65.93	7.41	20.04	3.55	1.65	2.41	13.09
GFN2-xTB	6.89	0.94	3.36	5.98	3.90	3.16	4.48	0.47	1.15	4.22
GFN1-xTB	5.82	0.96	3.92	5.55	5.50	3.73	3.94	0.56	4.96	4.19
DFTB3	11.13	1.24	3.23	7.85	6.07	8.01	6.60	0.75	2.59	5.76
DFTB/ChIMES	6.02	2.28	2.97	5.10	3.34	8.98	3.51	1.43	2.11	7.53
ODM2	17.57	1.34	5.20	5.66	4.37	6.22	12.06	0.92	4.62	4.07
PM6	17.39	1.10	4.24	8.12	7.90	5.97	10.65	0.80	5.72	3.37
PM6-D3H4X	18.31	0.96	4.59	6.23	7.72	7.04	11.68	0.59	5.58	3.92
PM7	16.13	0.99	3.44	6.19	6.32	7.19	9.83	0.64	4.04	3.55
AM1	18.64	1.92	3.68	8.37	3.85	5.77	14.14	2.09	3.94	4.18
MNDO/d	20.61	6.49	7.46	13.88	7.28	10.83	13.79	8.92	6.72	5.16
QD π -2	2.27	1.04	1.17	0.60	0.15	0.37	1.62	0.40	0.69	2.34

^a Only molecules containing the elements C, H, O, and N were included in the statistics. The “ANI” dataset refers to ANI-1⁴⁵ and ANI-2.⁴⁶ The COMP6⁵³ database consists of S66 $\times$ 8^{53,110,111} and COMP5.^{112–114} The relative energy (RE) dataset⁴² is a conglomeration of the following databases: Tautobase^{90,115} (TB), amino acid¹¹⁶ (AA), nucleic acid¹¹⁶ (NA), PA26¹¹⁷ (PA), HB375 $\times$ 10¹¹⁸ (HB), TAUT15¹¹⁷ (T15), and RegioSQM20¹¹⁹ (SQM).

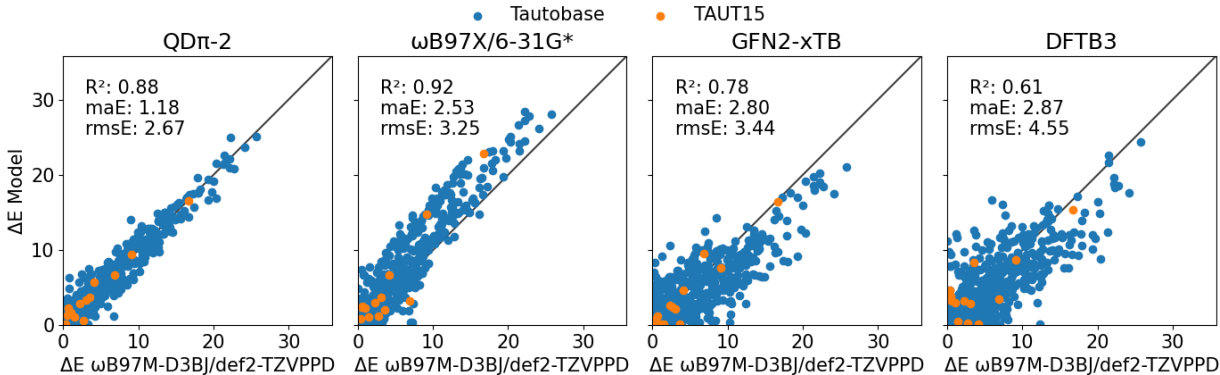


Figure 6: Comparison of tautomerization energies (kcal/mol). The reference values are ω B97M-D3BJ/def2-TZVPPD. The comparisons include molecules from the Tautobase^{90,115} and TAUT15¹¹⁷ databases.

Alchemical free energy simulation with QM/MM- Δ MLP “book-ending” corrections for ligand-protein binding

In this section we demonstrate that the new software infrastructure can be used to perform QM/MM- Δ MLP alchemical free energy calculations using an “indirect” approach. Specifically, the indirect approach is employed to calculate relative solvation free energies of the tautomers shown in Figure 2 with the QD π -2 model, whose parametrization was described in the previous section. Comparisons are made to DFTB3/3ob and GFN2-xTB QM/MM potentials and experiment, when available. Recall that the QD π -2 differs from the GFN2-xTB QM/MM model only by including a MLP correction to the QM/QM interactions; both models share the same QM/MM interactions.

Table 3 compares the experimental tautomeric free energies in solution to values obtained from free energy profiles evaluated from umbrella sampling. For clarity, these are the free energy differences between the solvated tautomeric forms; they are not relative solvation free energies. The DFTB3/3ob QM/MM potential yields the largest mean absolute error (maE). The GFN2-xTB QM/MM model reduces the error by approximately 50% (maE 3.3 kcal/mol). The QD π -2 QM/MM- Δ MLP reduces the error to 2.3 kcal/mol despite not having trained the neural network for the QM/MM interactions. These results show promise that the development of MLPs for the QM/MM interactions will lead to further improvement.

The MM $\rightarrow$ QM/MM- Δ MLP book-ending approach provides corrections to thermodynamic cycles where counterpoised transformations occur in two environments to obtain a relative free energy. When the environments are in the gas phase and aqueous solution, this implies a relative solvation free energy, whereas relative binding free energies calculate the free energies in solution and when complexed with a protein. We examine two strategies for evaluating MM $\rightarrow$ QM/MM- Δ MLP book-ending corrections. The first strategy analyzes equilibrium sampling produced by a λ -dependent potential energy function at a series of 21 uniformly spaced λ values. This strategy will be referred to as equilibrium book-ending (EBE). The second strategy analyzes nonequilibrium sampling with Jarzynski’s

Table 3: Tautomer reaction free energies (kcal/mol) in solution.^a

Reaction	Expt.	DFTB3		GFN2-xTB		QD π -2	
	ΔA	ΔA	$\sigma_{\Delta A}$	ΔA	$\sigma_{\Delta A}$	ΔA	$\sigma_{\Delta A}$
1A $\rightarrow$ 1B	-4.78	-12.93	0.12	-1.96	0.08	-2.93	0.12
2A $\rightarrow$ 2B	-6.10	-14.82	0.14	-6.53	0.06	-6.96	0.09
3A $\rightarrow$ 3B	-7.23	-14.81	0.13	-6.18	0.12	-8.59	0.06
4A $\rightarrow$ 4B	-2.30	-10.62	0.16	2.30	0.06	1.58	0.13
5A $\rightarrow$ 5B	-4.78	-9.75	0.08	-2.12	0.09	-2.13	0.09
6A $\rightarrow$ 6B	-9.28	-17.17	0.15	-8.75	0.11	-12.57	0.09
10D $\rightarrow$ 10C	-1.23	-6.88	0.17	5.12	0.07	0.95	0.08
12D $\rightarrow$ 12C	-1.77	-6.76	0.10	5.31	0.08	1.30	0.15
14D $\rightarrow$ 14C	0.27	-5.42	0.07	4.01	0.04	1.50	0.08
maE		6.88		3.25		2.26	
rmsE		7.03		3.97		2.46	

^a Experimental values taken from Ref. 90. $\sigma_{\Delta A}$ is the standard error of the mean from 4 trials. The summary statistics include the mean of absolute error (maE) and root-mean-square error (rmsE) relative to experiment.

equation,^{67,69–72} which we refer to as nonequilibrium book-ending (NEBE). These strategies are compared to solvation free energies obtained from umbrella sampling (US) performed in the aqueous and gas phase environments.

Table 4 compares the DFTB3/3ob, GFN2-xTB, and QD π -2 relative solvation free energies calculated from the US, EBE, and NEBE strategies. The US results are used here as a reference to compare the EBE and NEBE values, because the US calculations do not involve a MM reference potential. The mean absolute deviation of the EBE results is approximately 0.1 kcal/mol. The NEBE values have similar deviations for GFN2-xTB, slightly larger deviations for QD π -2 (maE 0.24 kcal/mol), and considerably higher deviations for DFTB3 (0.41 kcal/mol). The DFTB3 NEBE deviation is dominated by a 1.8 kcal/mol outlier (10D $\rightarrow$ 10C) that we are currently unable to rationalize. The NEBE deviations are related to the completeness of the equilibrium MM distribution and the non-equilibrium “switching time”. A more detailed study of how these parameters influence the precision of the calculation is forthcoming. Overall, the close agreement between the EBE and US results illustrate that the software infrastructure is in place for performing reliable QM/MM- Δ MLP alchemical

Table 4: Comparison of relative solvation free energies (kcal/mol).^a

Reaction	US		EBE		NEBE	
	$\Delta\Delta A$	$\sigma_{\Delta\Delta A}$	$\Delta\Delta A$	$\sigma_{\Delta\Delta A}$	ΔA	$\sigma_{\Delta\Delta A}$
DFTB3						
1A→1B	-15.27	0.16	-15.13	0.08	-15.37	0.09
2A→2B	-12.53	0.16	-12.66	0.09	-12.64	0.18
3A→3B	-12.01	0.16	-11.84	0.08	-12.23	0.14
4A→4B	-18.35	0.18	-18.41	0.10	-18.08	0.14
5A→5B	-12.80	0.16	-12.68	0.08	-12.48	0.12
6A→6B	-9.66	0.18	-9.48	0.08	-9.43	0.21
10D→10C	-19.43	0.19	-19.37	0.12	-18.65	0.29
12D→12C	-20.92	0.13	-20.76	0.18	-20.37	0.63
14D→14C	-14.15	0.10	-14.13	0.10	-13.65	0.20
maD			0.12		0.34	
rmsD			0.13		0.40	
GFN2-xTB						
1A→1B	-2.65	0.10	-2.71	0.06	-2.73	0.09
2A→2B	-2.52	0.07	-2.30	0.08	-2.42	0.10
3A→3B	-2.11	0.15	-2.25	0.07	-2.10	0.21
4A→4B	-3.17	0.08	-3.27	0.11	-2.90	0.21
5A→5B	-1.86	0.13	-1.89	0.10	-1.80	0.15
6A→6B	-0.95	0.16	-1.07	0.07	-0.92	0.16
10D→10C	-3.82	0.14	-3.72	0.09	-3.82	0.13
12D→12C	-3.91	0.14	-3.89	0.09	-3.76	0.12
14D→14C	-1.66	0.11	-1.67	0.07	-1.82	0.06
maD			0.09		0.10	
rmsD			0.11		0.13	
QD π -2						
1A→1B	-2.69	0.17	-2.73	0.06	-2.79	0.09
2A→2B	-2.30	0.14	-2.28	0.08	-2.29	0.17
3A→3B	-2.44	0.12	-2.21	0.09	-1.61	0.22
4A→4B	-3.20	0.15	-3.12	0.09	-2.96	0.13
5A→5B	-2.02	0.16	-1.89	0.09	-1.69	0.13
6A→6B	-0.91	0.16	-1.06	0.07	-1.05	0.14
10D→10C	-2.89	0.11	-2.83	0.09	-3.00	0.14
12D→12C	-3.31	0.17	-3.02	0.09	-2.97	0.10
14D→14C	-1.25	0.13	-1.24	0.06	-1.32	0.07
maD			0.11		0.24	
rmsD			0.14		0.33	

^a $\sigma_{\Delta\Delta A}$ is the standard error of the mean from 4 trials. The summary statistics include the mean absolute deviation (maD) and root-mean-square deviation (rmsD) relative to the values obtained from umbrella sampling.

free energy simulations using an indirect approach with book-ending corrections.

Conclusion

We report the implementation and testing of integrated software for performing free energy simulations with generalized hybrid QM/MM- Δ MLP force fields for enzyme mechanism and drug discovery applications. Specifically, we integrate Amber’s open-source MD and free energy program sander¹⁹ with the xtb²⁰ and DeePMD-kit^{21–23} to enable periodic boundary condition condensed phase simulations with rigorous treatment of the long-ranged electrostatics using QM/MM-Ewald methods.²⁴ We presented several example applications to demonstrate the capabilities of the software. The free energy surfaces of the SAR and MTR1 engineered RNA enzymes were calculated from umbrella sampling. The results indicate that the GFN2-xTB and DFTB3/3ob models perform reasonably well but neither model achieves quantitative agreement with *ab initio* PBE0/6-31G* QM/MM results. This suggests there is need to develop QM/MM- Δ MLPs to improve the accuracy of fast, semiempirical QM models to enable a broader range of practical applications. A preliminary QD π -2 force field for drug discovery was presented. The QD π -2 model supplements GFN2-xTB with a machine learning potential designed to improve internal QM/QM interactions. Comparisons were made to other models using a series of established databases, and we find that QD π -2 is superior to the first-generation QD π model based on DFTB3. Furthermore, the QD π -2 potential outperforms other machine-learning models and low-level *ab initio* DFT methods such as ω B97X/6-31G*. The new QD π -2 model was applied to alchemical free energy simulations using an indirect approach to calculate relative solvation free energies of small drug-like tautomers. It was demonstrated that the alchemical free energy infrastructure well-reproduced free energy differences obtained from umbrella sampling, which do not use a MM reference potential. Taken together, this work reports a significant advance in the creation of new integrated software infrastructure for the development and application

of new QM/MM- Δ MLP force fields for enzyme mechanism and drug discovery applications.

Supporting Information Available

Numerical validation tests of the sander/xtb Ewald implementation. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

Acknowledgments

The authors are grateful for the financial support provided by the National Institutes of Health (No. GM107485 and GM62248) and the National Science Foundation (CSSI Frameworks Grant No. 2209718). Computational resources were provided by the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program¹²⁰ (supercomputer Expanse at SDSC through allocation CHE190067), and by the Texas Advanced Computing Center (TACC) at the University of Texas at Austin (supercomputer Frontera¹²¹ through allocation CHE20002).

References

- (1) Chodera, J.; Mobley, D.; Shirts, M.; Dixon, R.; Branson, K.; Pande, V. Alchemical free energy methods for drug discovery: progress and challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150–160.
- (2) Beauchamp, K. A.; Lin, Y.-S.; Das, R.; Pande, V. S. Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements. *J. Chem. Theory Comput.* **2012**, *8*, 1409–1414.
- (3) Vangaveti, S.; Ranganathan, S. V.; Chen, A. A. Advances in RNA molecular dynamics: a simulator’s guide to RNA force fields. *Wiley Interdisciplinary Reviews: RNA* **2017**, *8*, 1396.

- (4) Wu, X.; Xu, L.-Y.; Li, E.-M.; Dong, G. Application of molecular dynamics simulation in biomedicine. *Chem. Biol. Drug Des.* **2022**, *99*, 789–800.
- (5) MacKerell, Jr., A. D. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* **2004**, *25*, 1584–1604.
- (6) Horn, H. W.; Swope, W. C.; Pitner, 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.
- (7) Joung, I. S.; Cheatham III, T. E. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *J. Phys. Chem. B* **2009**, *113*, 13279–13290.
- (8) Li, N.; Gao, Y.; Qiu, F.; Zhu, T. Benchmark Force Fields for the Molecular Dynamic Simulation of G-Quadruplexes. *Molecules* **2021**, *26*, 5379.
- (9) Boothroyd, S.; Madin, O. C.; Mobley, D. L.; Wang, L.-P.; Chodera, J. D.; Shirts, M. R. Improving Force Field Accuracy by Training against Condensed-Phase Mixture Properties. *J. Chem. Theory Comput.* **2022**, *18*, 3577–3592.
- (10) York, D. M. Modern Alchemical Free Energy Methods for Drug Discovery Explained. *ACS Phys. Chem. Au* **2023**, *3*, 478–491.
- (11) 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.
- (12) Giese, T. J.; Zeng, J.; Ekesan, S.; York, D. M. 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.* **2022**, *18*, 4304–4317.
- (13) Han, Y.; Wang, Z.; Wei, Z.; Liu, J.; Li, J. Machine learning builds full-QM precision protein force fields in seconds. *Brief. Bioinform.* **2021**, *22*, 1–10.

- (14) Zhang, Y.-J.; Khorshidi, A.; Kastlunger, G.; Peterson, A. A. The potential for machine learning in hybrid QM/MM calculations. *J. Chem. Phys.* **2018**, *148*, 241740.
- (15) Bösel, L.; Thürlmann, M.; Riniker, S. Machine Learning in QM/MM Molecular Dynamics Simulations of Condensed- Phase Systems. *J. Chem. Theory Comput.* **2021**, *17*, 2641–2658.
- (16) 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.
- (17) Lei, Y. K.; Yagi, K.; Sugita, Y. Learning QM/MM Potential using Equivariant Multi-scale Model. <https://doi.org/10.26434/chemrxiv-2024-80d93> (Accessed May 29, 2024).
- (18) Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Berryman, J. T.; Brozell, S. R.; Cerutti, D. S.; Cheatham, III, T. E.; Cisneros, G. A.; Cruzeiro, V. W. D. et al. AmberTools 2024. University of California, San Francisco, 2024.
- (19) Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzesh, N.; Giese, T. J.; Götz, A. W.; Gohlke, H. et al. AmberTools. *J. Chem. Inf. Model.* **2023**, *63*, 6183–6191.
- (20) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **2020**, *11*, e01493.
- (21) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184.
- (22) Liang, W.; Zeng, J.; York, D. M.; Zhang, L.; Wang, H. In *A Practical Guide to Recent*

Advances in Multiscale Modeling and Simulation of Biomolecules; Wang, Y., Zhou, R., Eds.; AIP Publishing, 2023; Chapter 6, pp 1–20.

- (23) Zeng, J.; Zhang, D.; Lu, D.; Mo, P.; Li, Z.; Chen, Y.; Rynik, M.; Huang, L.; Li, Z.; Shi, S. et al. DeePMD-kit v2: A software package for deep potential models. *J. Chem. Phys.* **2023**, *159*, 054801.
- (24) 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.
- (25) Wilson, C.; Szostak, J. In vitro evolution of a self-alkylating ribozyme. *Nature* **1995**, *374*, 777–782.
- (26) Krochmal, D.; Shao, Y.; Li, N.-S.; DasGupta, S.; Shelke, S. A.; Koirala, D.; Piccirilli, J. A. Structural basis for substrate binding and catalysis by a self-alkylating ribozyme. *Nat. Chem. Biol.* **2022**, *18*, 376–384.
- (27) Deng, J.; Wilson, T. J.; Wang, J.; Peng, X.; Li, M.; Lin, X.; Liao, W.; Lilley, D. M. J.; Huang, L. Structure and mechanism of a methyltransferase ribozyme. *Nat. Chem. Biol.* **2022**, *18*, 556–564.
- (28) Scheitl, C. P. M.; Mieczkowski, M.; Schindelin, H.; Höbartner, C. Structure and mechanism of the methyltransferase ribozyme MTR1. *Nat. Chem. Biol.* **2022**, *18*, 547–555.
- (29) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
- (30) Scheitl, C. P. M.; Ghaem Maghami, M.; Lenz, A.-K.; Höbartner, C. Site-specific RNA methylation by a methyltransferase ribozyme. *Nature* **2020**, *587*, 663–667.

- (31) Giese, T. J.; York, D. M. Development of a Robust Indirect Approach for MM $\rightarrow$ QM Free Energy Calculations That Combines Force-Matched Reference Potential and Bennett’s Acceptance Ratio Methods. *J. Chem. Theory Comput.* **2019**, *15*, 5543–5562.
- (32) Grimme, S.; Bannwarth, C.; Shushkov, P. 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.* **2017**, *13*, 1989–2009.
- (33) Grimme, S.; Bannwarth, C.; Caldeweyher, E.; Pisarek, J.; Hansen, A. A general intermolecular force field based on tight-binding quantum chemical calculations. *J. Chem. Phys.* **2017**, *147*, 161708.
- (34) Giese, T. J.; Panteva, M. T.; Chen, H.; York, D. M. Multipolar Ewald methods, 1: Theory, accuracy, and performance. *J. Chem. Theory Comput.* **2015**, *11*, 436–450.
- (35) 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.
- (36) 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.
- (37) Dral, P. O.; Zubatiuk, T.; Xue, B.-X. In *Quantum Chemistry in the Age of Machine Learning*; Dral, P. O., Ed.; Elsevier, 2022; Chapter 21, pp 491–507.
- (38) Liu, Y.; Li, J. Permutation-Invariant-Polynomial Neural-Network-Based Δ -Machine Learning Approach: A Case for the HO₂ Self-Reaction and Its Dynamics Study. *J. Phys. Chem. Lett.* **2022**, *13*, 4729–4738.

- (39) Ruth, M.; Gerbig, D.; Schreiner, P. R. Machine Learning of Coupled Cluster (T)-Energy Corrections via Delta (Δ)-Learning. *J. Chem. Theory Comput.* **2022**, *18*, 4846–4855.
- (40) Unzueta, P. A.; Greenwell, C. S.; Beran, G. J. O. Predicting Density Functional Theory-Quality Nuclear Magnetic Resonance Chemical Shifts via Δ -Machine Learning. *J. Chem. Theory Comput.* **2021**, *17*, 826–840.
- (41) Zeng, J.; Giese, T. J.; Ekesan, Ş.; York, D. M. Development of Range-Corrected Deep Learning Potentials for Fast, Accurate Quantum Mechanical/Molecular Mechanical Simulations of Chemical Reactions in Solution. *J. Chem. Theory Comput.* **2021**, *17*, 6993–7009.
- (42) Zeng, J.; Tao, Y.; Giese, T. J.; York, D. M. QD π : A Quantum Deep Potential Interaction Model for Drug Discovery. *J. Chem. Theory Comput.* **2023**, *19*, 1261–1275.
- (43) Zhang, D.; Bi, H.; Dai, F.-Z.; Jiang, W.; Liu, X.; Zhang, L.; Wang, H. Pretraining of attention-based deep learning potential model for molecular simulation. *npj Comput Mater* **2024**, *10*, 94.
- (44) Eastman, P.; Behara, P. K.; Dotson, D. L.; Galvelis, R.; Herr, J. E.; Horton, J. T.; Mao, Y.; Chodera, J. D.; Pritchard, B. P.; Wang, Y. et al. SPICE, A Dataset of Drug-like Molecules and Peptides for Training Machine Learning Potentials. *Sci. Data* **2023**, *10*, 11.
- (45) Smith, J. S.; Zubatyuk, R.; Nebgen, B.; Lubbers, N.; Barros, K.; Roitberg, A. E.; Isayev, O.; Tretiak, S. The ANI-1ccx and ANI-1x data sets, coupled-cluster and density functional theory properties for molecules. *Sci Data* **2020**, *7*, 134.
- (46) Devereux, C.; Smith, J. S.; Huddleston, K. K.; Barros, K.; Zubatyuk, R.; Isayev, O.; Roitberg, A. E. Extending the Applicability of the ANI Deep Learning Molecular Potential to Sulfur and Halogens. *J. Chem. Theory Comput.* **2020**, *16*, 4192–4202.

- (47) Axelrod, S.; Gómez-Bombarelli, R. GEOM, energy-annotated molecular conformations for property prediction and molecular generation. *Sci. Data* **2022**, *9*, 185.
- (48) Mobley, D. L.; Guthrie, J. P. FreeSolv: a database of experimental and calculated hydration free energies, with input files. *J. Comput. Aid. Mol. Des.* **2014**, *28*, 711–720.
- (49) Zhang, Y.; Wang, H.; Chen, W.; Zeng, J.; Zhang, L.; Han, W.; E, W. DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun.* **2020**, *253*, 107206.
- (50) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.
- (51) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A. et al. P<scp>SI4</scp> 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.
- (52) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.
- (53) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. Less is more: Sampling chemical space with active learning. *J. Chem. Phys.* **2018**, *148*, 241733–241743.
- (54) Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.
- (55) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the self-consistent-charge

- density-functional tight-binding method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.
- (56) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.
- (57) Hourahine, B.; Aradi, B.; Blum, V.; Bonafe, F.; Buccheri, A.; Camacho, C.; Cevallos, C.; Deshayre, M. Y.; Dumitrica, T.; Dominguez, A. et al. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **2020**, *152*, 124101.
- (58) Zeng, J.; Tao, Y.; Giese, T. J.; York, D. M. Modern semiempirical electronic structure methods and machine learning potentials for drug discovery: Conformers, tautomers, and protonation states. *J. Chem. Phys.* **2023**, *158*, 124110.
- (59) 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.
- (60) Tsai, H.-C.; Tao, Y.; Lee, T.-S.; Merz, K. M.; York, D. M. Validation of Free Energy Methods in AMBER. *J. Chem. Inf. Model.* **2020**, *60*, 5296–5300.
- (61) Tsai, H.-C.; Lee, T.-S.; Ganguly, A.; Giese, T. J.; Ebert, M. C.; Labute, P.; Merz Jr, K. M.; York, D. M. AMBER Free Energy Tools: A New Framework for the Design of Optimized Alchemical Transformation Pathways. *J. Chem. Theory Comput.* **2023**, *19*, 640–658.
- (62) Lee, T.-S.; Lin, Z.; Allen, B. K.; Lin, C.; Radak, B. K.; Tao, Y.; Tsai, H.-C.; Sherman, W.; York, D. M. Improved Alchemical Free Energy Calculations with Optimized Smoothstep Softcore Potentials. *J. Chem. Theory Comput.* **2020**, *16*, 5512–5525.

- (63) Lee, T.-S.; Tsai, H.-C.; Ganguly, A.; Giese, T. J.; York, D. M. In *Robust, Efficient and Automated Methods for Accurate Prediction of Protein-Ligand Binding Affinities in AMBER Drug Discovery Boost*; Armacost, K. A., Thompson, D. C., Eds.; ACS Symposium Series; 2021; Vol. 1397; pp 161–204.
- (64) Lee, T.-S.; Tsai, H.-C.; Ganguly, A.; York, D. M. ACES: Optimized Alchemically Enhanced Sampling. *J. Chem. Theory Comput.* **2023**, *19*, 472–487.
- (65) König, G.; Hudson, P. S.; Boresch, S.; Woodcock, H. L. Multiscale Free Energy Simulations: An Efficient Method for Connecting Classical MD Simulations to QM or QM/MM Free Energies Using Non-Boltzmann Bennett Reweighting Schemes. *J. Chem. Theory Comput.* **2014**, *10*, 1406–1419.
- (66) König, G.; Mei, Y.; Pickard 4th, F. C.; Simmonett, A. C.; Miller, B. T.; Herbert, J. M.; Woodcock, H. L.; Brooks, B. R.; Shao, Y. Computation of Hydration Free Energies Using the Multiple Environment Single System Quantum Mechanical/Molecular Mechanical Method. *J. Chem. Theory Comput.* **2016**, *12*, 332–44.
- (67) Kearns, F. L.; Hudson, P. S.; Boresch, S.; Woodcock, H. L. Methods for Efficiently and Accurately Computing Quantum Mechanical Free Energies for Enzyme Catalysis. *Methods Enzymol.* **2016**, *577*, 75–104.
- (68) Hudson, P. S.; White, J. K.; Kearns, F. L.; Hodoscek, M.; Boresch, S.; Woodcock, H. L. Efficiently computing pathway free energies: New approaches based on chain-of-replica and Non-Boltzmann Bennett reweighting schemes. *Biochim. Biophys. Acta* **2015**, *1850*, 944–953.
- (69) Hudson, P. S.; Woodcock, H. L.; Boresch, S. Use of Nonequilibrium Work Methods to Compute Free Energy Differences Between Molecular Mechanical and Quantum Mechanical Representations of Molecular Systems. *J. Phys. Chem. Lett.* **2015**, *6*, 4850–4856.

- (70) Kearns, F. L.; Hudson, P. S.; Woodcock, H. L.; Boresch, S. Computing converged free energy differences between levels of theory via nonequilibrium work methods: Challenges and opportunities. *J. Comput. Chem.* **2017**, *38*, 1376–1388.
- (71) Boresch, S.; Woodcock, H. L. Convergence of single-step free energy perturbation. *Mol. Phys.* **2017**, *115*, 1200–1213.
- (72) Schöller, A.; Kearns, F.; Woodcock, H. L.; Boresch, S. Optimizing the Calculation of Free Energy Differences in Nonequilibrium Work SQM/MM Switching Simulations. *J. Phys. Chem. B* **2022**, *126*, 2798–2811.
- (73) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (74) Jarzynski, C. Equilibrium free-energy differences from nonequilibrium measurements: a master-equation approach. *Phys. Rev. E* **1997**, *56*, 5018–5035.
- (75) Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.
- (76) Crooks, G. E. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E* **2000**, *61*, 2361–2366.
- (77) Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham III, T. E.; Jurečka, P. Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886–2902.
- (78) Joung, I. S.; Cheatham III, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.

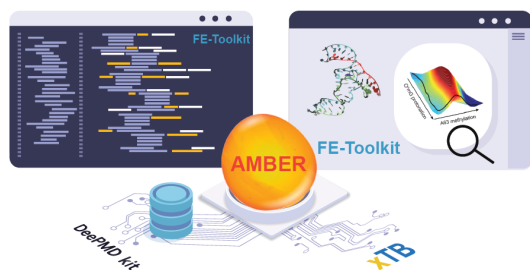
- (79) Figueirido, F.; Del Buono, G. S.; Levy, R. M. On finite-size effects in computer simulations using the Ewald potential. *J. Chem. Phys.* **1995**, *103*, 6133–6142.
- (80) de Leeuw, S. W.; Perram, J. W.; Smith, E. R. Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants. *Proc. R. Soc. London, Ser. A* **1980**, *373*, 27–56.
- (81) 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.
- (82) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (83) Tan, Z.; Gallicchio, E.; Lapelosa, M.; Levy, R. M. Theory of binless multi-state free energy estimation with applications to protein-ligand binding. *J. Chem. Phys.* **2012**, *136*, 144102.
- (84) Zhang, B. W.; Xia, J.; Tan, Z.; Levy, R. M. A Stochastic Solution to the Unbinned WHAM Equations. *J. Phys. Chem. Lett.* **2015**, *6*, 3834–3840.
- (85) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105.
- (86) Giese, T. J.; York, D. M. FE-ToolKit: The free energy analysis toolkit. <https://gitlab.com/RutgersLBSR/fe-toolkit> (Accessed May 29, 2024).
- (87) Giese, T. J.; Ekesan, Ş.; McCarthy, E.; Tao, Y.; York, D. M. Surface-Accelerated String Method for Locating Minimum Free Energy Paths. *J. Chem. Theory Comput.* **2024**, *20*, 2058–2073.
- (88) McCarthy, E.; Ekesan, Ş.; Giese, T. J.; Wilson, T. J.; Deng, J.; Huang, L.; Lilley, D. M. J.; York, D. M. Catalytic mechanism and pH dependence of a methyltransferase

- ribozyme (MTR1) from computational enzymology. *Nucleic Acids Res.* **2023**, *51*, 4508–4518.
- (89) Sitzmann, M.; Ihlenfeldt, W.-D.; Nicklaus, M. C. Tautomerism in large databases. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 521–51.
- (90) Wahl, O.; Sander, T. Tautobase: An Open Tautomer Database. *J. Chem. Inf. Model.* **2020**, *60*, 1085–1089.
- (91) Dhaked, D. K.; Guasch, L.; Nicklaus, M. C. Tautomer Database: A Comprehensive Resource for Tautomerism Analyses. *J. Chem. Inf. Model.* **2020**, *60*, 1090–1100.
- (92) He, X.; Man, V. H.; Yang, W.; Lee, T.-S.; Wang, J. A fast and high-quality charge model for the next generation general AMBER force field. *J. Chem. Phys.* **2020**, *153*, 114502.
- (93) Walker, R. C.; Crowley, M. F.; Case, D. A. The implementation of a fast and accurate QM/MM potential method in Amber. *J. Comput. Chem.* **2008**, *29*, 1019–1031.
- (94) Seabra, G.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package. *J. Phys. Chem. A* **2007**, *111*, 5655–5664.
- (95) Nishimoto, K.; Mataga, N. Electronic structure and spectra of some nitrogen heterocycles. *Z. Phys. Chem.* **1957**, *12*, 335–338.
- (96) Ohno, K. Some remarks on the Pariser-Parr-Pople method. *Theor. Chim. Act.* **1964**, *2*, 219–227.
- (97) Klopman, G. A semiempirical treatment of molecular structures, II. Molecular terms and application to diatomic molecules. *J. Am. Chem. Soc.* **1964**, *86*, 4550–4557.
- (98) Ma, L.; Liu, J. Catalytic Nucleic Acids: Biochemistry, Chemical Biology, Biosensors, and Nanotechnology. *iScience* **2020**, *23*, 100815.

- (99) Balke, D.; Hieronymus, R.; Müller, S. *Catalytically Active Nucleic Acids*; 2017; Vol. 170; Chapter 2, pp 21–35.
- (100) Müller, S.; Appel, B.; Balke, D.; Hieronymus, R.; Nübel, C. Thirty-five years of research into ribozymes and nucleic acid catalysis: where do we stand today? *F1000 Res.* **2016**, *5*, 1511.
- (101) Yan, J.; Ran, M.; Shen, X.; Zhang, H. Therapeutic DNazymes: From Structure Design to Clinical Applications. *Adv. Mater.* **2023**, *35*, 2300374.
- (102) Chen, X.; Li, N.; Ellington, A. D. Ribozyme catalysis of metabolism in the RNA world. *Chem. Biodivers.* **2007**, *4*, 633–655.
- (103) Piletic, I. R.; Edney, E. O.; Bartolotti, L. J. A computational study of acid catalyzed aerosol reactions of atmospherically relevant epoxides. *Phys. Chem. Chem. Phys.* **2013**, *15*, 18065–18076.
- (104) Giese, T. J.; Zeng, J.; York, D. M. Multireference Generalization of the Weighted Thermodynamic Perturbation Method. *J. Phys. Chem. A* **2022**, *126*, 8519–8533.
- (105) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. <https://doi.org/10.48550/arXiv.1412.6980> (Accessed May 29, 2024).
- (106) Martin, Y. C. Experimental and pK prediction aspects of tautomerism of drug-like molecules. *Drug Discov. Today. Technol.* **2018**, *27*, 59–64.
- (107) Milletti, F.; Vulpetti, A. Tautomer Preference in PDB Complexes and its Impact on Structure-Based Drug Discovery. *J. Chem. Inf. Model.* **2010**, *50*, 1062–1074.
- (108) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (109) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L. et al. Psi4: an

- open-source ab initio electronic structure program. *WIREs Comput. Mol. Sci.* **2012**, *2*, 556–565.
- (110) Goerigk, L.; Kruse, H.; Grimme, S. Benchmarking Density Functional Methods against the S66 and S66x8 Datasets for Non-Covalent Interactions. *Chem. Phys. Chem.* **2011**, *12*, 3421–3433.
- (111) Brauer, B.; Kesharwani, M. K.; Kozuch, S.; Martin, J. M. L. The S66x8 benchmark for noncovalent interactions revisited: explicitly correlated ab initio methods and density functional theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20905–20925.
- (112) Fink, T.; Bruggesser, H.; Reymond, J.-L. Virtual Exploration of the Small-Molecule Chemical Universe below 160 Daltons. *Angew. Chem. Int. Ed.* **2005**, *44*, 1504–1508.
- (113) Blum, L. C.; Reymond, J.-L. 970 Million Druglike Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13. *J. Am. Chem. Soc.* **2009**, *131*, 8732–8733.
- (114) Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A. C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V. et al. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.* **2014**, *42*, 1091–1097.
- (115) Wieder, M.; Fass, J.; Chodera, J. D. Fitting quantum machine learning potentials to experimental free energy data: predicting tautomer ratios in solution. *Chem. Sci.* **2021**, *12*, 11364–11381.
- (116) Moser, A.; Range, K.; York, D. M. Accurate Proton Affinity and Gas-Phase Basicity Values for Molecules Important in Biocatalysis. *J. Phys. Chem. B* **2010**, *114*, 13911–13921.
- (117) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main

- group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.
- (118) Řezáč, J. Non-Covalent Interactions Atlas Benchmark Data Sets: Hydrogen Bonding. *J. Chem. Theory Comput.* **2020**, *16*, 2355–2368.
- (119) Ree, N.; Göller, A. H.; Jensen, J. H. RegioSQM20: improved prediction of the regioselectivity of electrophilic aromatic substitutions. *J. Cheminform.* **2021**, *13*, 10.
- (120) Boerner, T. J.; Deems, S.; Furlani, T. R.; Knuth, S. L.; Towns, J. In *PEARC'23: Practice and Experience in Advanced Research Computing*; Romanella, R. S. A., Knuth, S., Hackworth, K., Pummill, J., Eds.; ACCESS: Advancing Innovation: NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services & Supports; Association for Computing Machinery, 2023; pp 173–176.
- (121) Stanzione, D.; West, J.; Evans, R. T.; Minyard, T.; Ghattas, O.; Panda, D. K. *PEARC'20: Practice and Experience in Advanced Research Computing*; Frontera: The Evolution of Leadership Computing at the National Science Foundation; Association for Computing Machinery, 2020; pp 106–111.



TOC Graphic