Benchmarking Force Field and the ANI Neural Network Potentials for the Torsional Potential Energy Surface of Biaryl Drug Fragments

Shae-Lynn J. Lahey,[†] Tu Nguyen Thien Phuc,[†] and Christopher N. Rowley[*,†,‡]

[†]*Department of Chemistry, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada*

[‡]*A shared footnote*

E-mail: crowley@mun.ca

Phone: +1 (709) 864-2229

Abstract

Many drug molecules contain biaryl fragments, resulting in a torsional barrier corresponding to rotation around the bond linking the aryls. The potential energy surfaces of these torsions vary significantly due to steric and electronic effects, ultimately affecting the relative stability of the molecular conformations in the protein-bound and solution states. Simulations of protein–ligand binding require accurate computational models to represent the intramolecular interactions to provide accurate predictions of the structure and dynamics of binding. In this paper, we compare four force fields (Generalized AMBER Force Field (GAFF), Open Force Field (OpenFF), CHARMM General Force Field (CGenFF), Optimized Potentials for Liquid Simulations (OPLS)) and two neural network potentials (ANI-2x, ANI-1ccx) in their ability to predict the torsional potential energy surfaces of 88 biaryls extracted from drug fragments. The mean of the absolute deviation over the full PES (MADF) and the mean absolute

deviation of the torsion rotational barrier height (MADB) relative to high-level ab initio reference data was used as a measure of accuracy. In comparison to high-level ab-initio data, ANI-1ccx was most accurate for predicting the barrier height (MADF: 0.5 kcal/mol, MADB: 0.8 kcal/mol), followed closely by ANI-2x (MADF: 0.5 kcal/mol, MADB: 1.0 kcal/mol), then CGenFF (MADF: 0.8 kcal/mol, MADB: 1.3 kcal/mol), OpenFF (MADF: 1.5 kcal/mol, MADB: 1.4 kcal/mol), GAFF (MADF: 1.2 kcal/mol, MADB: 2.6 kcal/mol), and finally OPLS (MADF: 1.5 kcal/mol, MADB: 2.8 kcal/mol). Significantly, the NNPs are systematically more accurate and more reliable than any of the force fields. As a practical example, the neural network potential/molecular mechanics (NNP/MM) method was used to simulate the isomerization of ozanimod, a drug used for multiple sclerosis. Multi-nanosecond molecular dynamics (MD) simulations in an explicit aqueous solvent were performed, as well as umbrella sampling and adaptive biasing force enhanced sampling techniques. These theories predicted a rate of isomerization of 4.30×10^{-1} ns^{-1}, which is consistent with direct MD simulations.

1 Introduction

Small-molecule ligands of biological molecules can hold a range of geometries, both in solution and in their bound state. The loss of conformational freedom and distortion of molecules in their bound state can result in a significant thermodynamic penalty opposing ligand binding.[1–3] Simulations of protein–ligand binding require accurate methods to calculate the intramolecular interactions of these ligands to predict the relative stability of these conformations.

Many natural products and drug molecules contain biaryl motifs and rotation around the bonds connecting them that introduces a torsional degree of freedom in these molecules. The potential energy surface for the rotation around these bonds varies due to conjugation, steric interactions, intramolecular hydrogen bonding, and electron repulsion. These effects determine the equilibrium conformations held by the molecules and the rates of conforma-

tional isomerization. Accurate computational models for these torsional potential energy surfaces is essential for modeling conformational dynamics and protein–ligand binding.

Although high-level ab initio methods (e.g., MP2 or CCSD) provide accurate predictions of the stability of molecular geometry, simulations of protein–ligand binding commonly require the evaluation of millions or billions of configurations, making these methods impractical to be used directly. Instead, molecular mechanical (MM) force fields are defined that approximate the intramolecular interactions in terms of simple mathematical functions.[4] Torsional potentials are generally defined as a sum of cosine functions with a variety of periods, offsets, and amplitudes,

$$\mathcal{V}_{\text{torsion}}(\theta) = \sum_i k_i \left(1 + \cos\left(n_i\theta + \phi_i\right)\right) \tag{1}$$

Additionally, intramolecular non-bonded interactions can significantly affect torsional potential energy surfaces. In some biaryls, intramolecular non-bonded interactions between atoms beyond the 1,4 interactions also significantly affect torsional potential energy surfaces due to steric or electrostatic interactions.[5]

In conventional molecular mechanical force fields, the parameters for the torsional and non-bonded potentials are defined by assigning each atom an "atom type" based on its element and chemical environment.[6] Appropriate parameters must then be defined for all relevant permutations of atom types that define a torsional interaction. These present significant challenges because accurate coverage of the full variety of chemical bonding motifs present in drug-like molecules can require hundreds or thousands of parameters to be defined.[7]

Innovations have been introduced to simplify this process. New methods have been developed to determine parameters for the possible permutations of atom types automatically. Where a force field lacks a specific term, this parameter can be fit to an ab initio potential energy surface. The SMIRKS Native Open Force Field (SMIRNOFF) format assigns force field parameters by searching for chemical patterns.[8] These methods still require the defini-

tion of extensive sets of parameters and follow most of the standard approximations inherent to conventional force fields.

Recently, general-purpose neural network potentials (NNPs) have emerged as an alternative to molecular mechanical force fields.[9] In these methods, neural networks are trained to predict the potential energy for an arbitrary molecular geometry, with the promise of providing ab initio quality results. The ANI neural network potentials developed by Smith, Isayev, and Roitberg are notably successful NNPs.[10–12] This model defines a set of "symmetry functions" for an atom that encodes its immediate chemical environment.[13] A neural network is trained to reproduce ab initio electronic energies for a training set of molecules from these symmetry functions. The ANI-2x potential was trained to reproduce density functional theory (DFT) calculated energies (ωB97X-D/6-31G*) of 5 million different molecular geometries for compounds containing elements C, N, O, H, F, S, and Cl.[12] The ANI-1ccx NNP was developed using transfer learning, where the inner layers of the ANI-1x model were transferred while the input and output layers were trained to CCSD(T)*/CBS data, although it is limited to the elements C, N, O, and H.[11] Importantly, these models have shown remarkable transferability; they provide accurate predictions of molecules that are not present in their training sets. Further, they avoid the standard force field approximations where intramolecular interactions are cast into harmonic, cosine, Coulombic, and Lennard-Jones potentials.

Replacing MM models with NNP models in simulations of protein–ligand interactions could provide ab initio accuracy at a similar computational cost to molecular mechanical models while avoiding the parameterization of individual ligands. Our group recently showed that the ANI-1ccx NNP is effective for representing the intramolecular terms of protein-bound ligands when embedded into a conventional MM model (torsional).[14] This method, termed NNP/MM (a.k.a., ML/MM), allows NNPs to be used to represent the intramolecular terms of a ligand in protein–ligand binding simulations where the solvent and protein are represented using a molecular mechanical (MM) model. The NNP/MM model provided

distinct conformational components of the absolute binding energy in comparison to a pure MM model, notably in the case of erlotinib bound to the Epidermal Growth Factor Receptor, where the CGenFF force field spuriously predicted a large Gibbs energy penalty for the ligand to adopt a bound conformation. Rufa and coworkers later showed that the accuracy of relative alchemical free energy perturbation of protein–ligand binding calculations was improved by perturbing between ANI-2x/OpenFF NNP/MM models.[15]

NNP/MM has also served as a more effective ligand model in molecular dynamics flexible fitting (MDFF) refinement of CryoEM structures.[16] Similarly, a machine learning-based intramolecular potential was used by Cole et al. to simulate the conformational dynamics of 3-(benzyloxy)pyridine-2-amine in protein-bound and solution states.[17]

Biaryl torsions present a challenging and important test of the force field and NNP models for the conformations of drug-like molecules. The π-systems of the aryl rings affect the torsional potential energy surface in a non-local minimum and the interplay of several types of intramolecular interactions. Jorgensen and coworkers published a test set of biaryls present in drug molecules and drug candidates, which provides an extensive, diverse, and relevant test set for assessing the accuracy of computational methods for predicting biaryl torsional potential energy surfaces.[18]

In this paper, we compare the performance of four molecular mechanical models (GAFF, CGenFF, OPLS, and OpenFF) and two NNPs (ANI-2x and ANI-1ccx) for the prediction of biaryl torsional potential energy surfaces calculated using high-level ab initio method (CCSD(T1*)/CBS).

2 Computational Methods

2.1 Test Set

The test set of biaryls used here is largely the same as the biaryl torsional test set developed by Dahlgren et al.[18] This test set was generated by extracting biaryl fragments from drug and

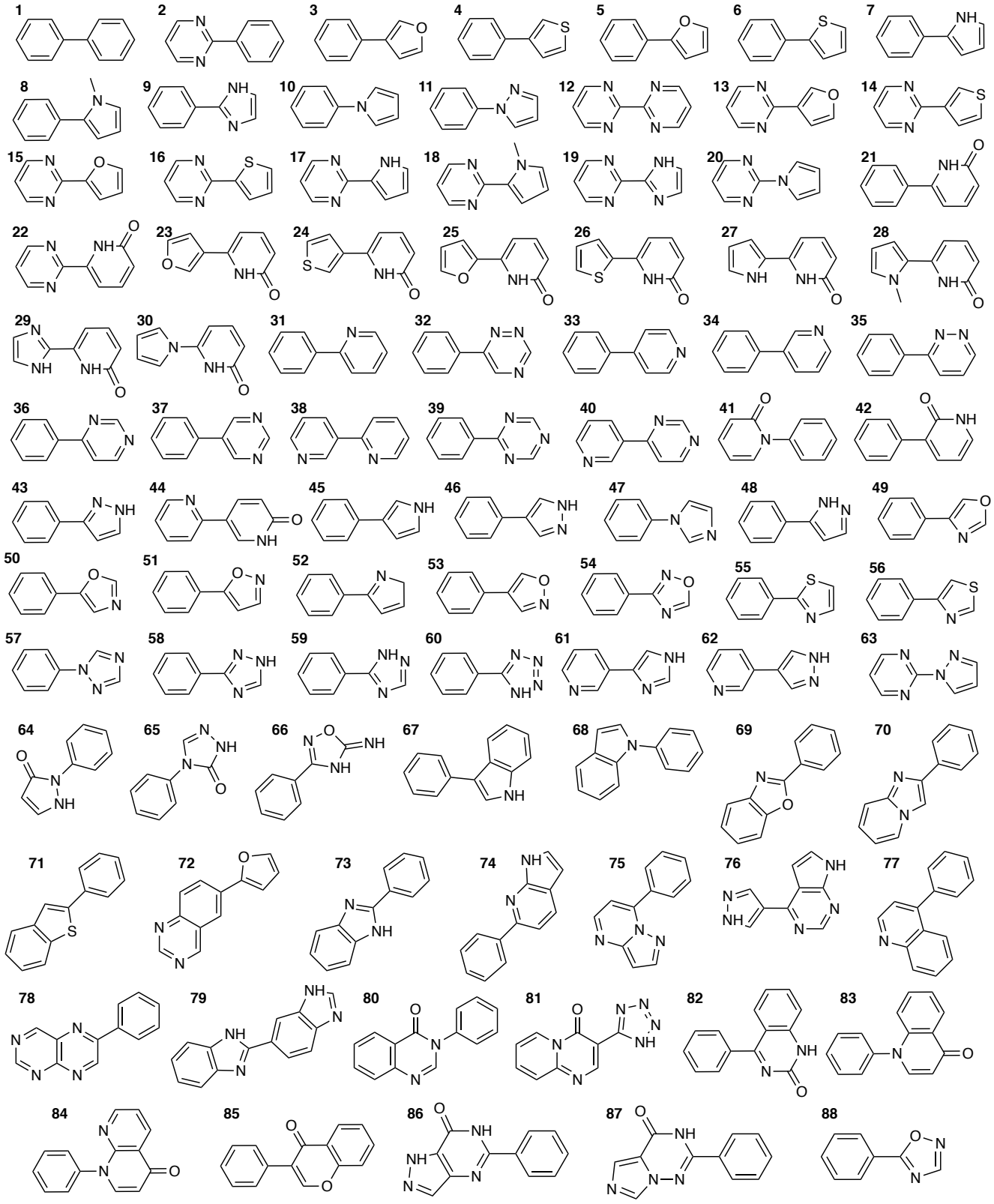

Figure 1: Test set of biaryl fragments.

drug-like molecules. Because some of the methods used here are not designed for use with charged compounds (e.g., the neural network potentials), charged compounds were excluded from the test set. Also, 5-phenyl-1,2,4-oxadiazole, a fragment of the drug ozanimod, was added to our test set. The structures of the molecules in the test set and their associated numbering are illustrated in Figure 1. The structures, topology, and parameter files of this test set are available on our GitHub repository.[19]

2.1.1 Molecular Mechanical Parameterization and Calculations

The CGenFF, GAFF, and OPLS potential energy surfaces were calculated using relaxed scans using CHARMM (i.e., the torsional degree of freedom was restrained and the remaining degrees of freedom were energy-minimized). These calculations were performed using CHARMM 41b2.[20]

CGenFF calculations employed the CGenFF force field version 2.2.0.[21,22] The charge, topology, and parameters were assigned using the CGenFF parameterization server. 1,4 non-bonded interactions calculated using the CGenFF force field were not scaled.[23,24]

The GAFF parameters were assigned using AmberTools.[25] Atomic charges in the GAFF model were calculated using the Restrained Electrostatic Potential method (RESP) using Hartree–Fock (HF) with the 6-31G* basis set as the target quantum mechanical (QM) electrostatic potential.[26] 1,4 nonbonded interactions in the GAFF force field calculations were scaled by a factor of 0.833.

The OPLS[27] topology and parameters were generated using the LigParGen server.[28] The CM1A-LBCC charge model was used.[29] 1,4 nonbonded interactions in the OPLS force field were scaled by a factor of 0.5. The geometric combination rule was used for the σ Lennard-Jones parameters.

The OpenFF charges and parameters were generated using the SMIRNOFF Open Force Field version 1.1.1 with the code name "Parsley".[8] The relaxed potential scans were performed in OpenMM package version 7.4.1[30] with external harmonic restraint on the interest

dihedral angle. The harmonic strength was set to ensure the torsional angle variance was less than 0.02°.

2.2 ANI Potential Energy Surfaces

The potential energy surfaces for the ANI-2x[12] and ANI-1ccx[11] NNPs were calculated using TorchANI[31] interfaced with the External feature of Gaussian 09[32] through a python script. This script is available for download from our GitHub repository. Each surface was calculated from complete scans in the forward and reverse directions for rotation around this torsion, where the minimum of the energies from each scan was used to construct the PES.

2.3 QM Potential Energy Surfaces

The QM potential energy surfaces were calculated using a relaxed scan of the biaryl torsional degree of freedom using the RIMP2/def2-TZVP level of theory.[33–35] ORCA 4.2.1[36] was used to calculate the single point potential energy of these configurations using the composite CCSD(T)*/CBS method described by Smith et al.[11] The iterative triples method, DLNPO-CCSD(T1),[37] was used because the DLNPO-CCSD(T) torsional potential energy surfaces of some molecules were discontinuous due to differences in the conventional triples correction energy.

2.4 NNP/MM MD Simulations

An NNP/MM simulation was performed of ozanimod (5-[3-[(1S)-1-(2-hydroxyethylamino)-2,3-dihydro-1H-inden-4-yl]-1,2,4-oxadiazol-5-yl]-2-propan-2-yloxybenzonitrile) in an explicit aqueous solvent. In this method, the intramolecular potential energy of the ligand ($\mathcal{V}_{NNP}(\mathbf{r}_{NNP})$) is calculated using the ANI-1ccx NNP, while the potential energy of the solvent is represented using a conventional molecular mechanical model ($\mathcal{V}_{NNP}(\mathbf{r}_{NNP})$). The potential energy of the system are the sum of these two components and an additional term corresponding to

the interaction of the NNP and MM atoms. (Eqn. 3).

$$\mathcal{V}(\mathbf{r}) = \mathcal{V}_{MM}\left(\mathbf{r}_{MM}\right) + \mathcal{V}_{NNP}\left(\mathbf{r}_{NNP}\right) + \mathcal{V}_{NNP/MM}\left(\mathbf{r}\right) \tag{2}$$

The interactions between the atoms represented using the NNP and atoms represented using MM are calculated using Lennard-Jones and Coulombic potentials (Eqn. **??**).

$$\mathcal{V}_{NNP/MM}\left(\mathbf{r}\right) = \sum_{i}^{MM}\sum_{j}^{NNP}\frac{q_i q_j}{4\pi\epsilon r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{3}$$

In these simulations, the solvent–solute Lennard-Jones parameters (σ and ϵ)[25] were generated using the Lorentz–Berthelot combination rules using the GAFF parameters for the solute and the TIP3P-FB parameters for the water molecules. The partial atomic charges (q) of the solute were calculated using the RESP method.[26]

The simulations were performed using NAMD 2.13[38] interfaced with TorchANI[31] using the NAMD-ANI interface script.[39] The total number of water molecules was 2158. The dimensions of the periodic simulation cell were 48.7 Å ×38.5Å ×34.8 Å. The ozanimod molecule was represented using the ANI-1ccx NNP and the solvent was represented using the TIP3P-FB water model.[40] The GAFF Lennard-Jones parameters were used for the solute–solvent non-bonded interactions.[25] Intermolecular electrostatic interactions were calculated based on RESP charges assigned to the atoms of the ozanimod.[26] In the NNP/MM framework, the intramolecular interactions of the ligand are calculated using only the NNP.[14] A 2 fs timestep was used. Bonds containing hydrogen atoms were constrained using the SHAKE algorithm. In these simulations, the temperature was coupled to a 298.15 K bath using a Lowe–Anderson thermostat.[41]

Calculation of the potential of mean force (PMF) for the rotation of the N-C-C-C biaryl torsion was performed using the adaptive biasing force (ABF)[42,43] and umbrella sampling.[44,45] Umbrella sampling simulations were performed on the N-C-C-C biaryl torsional coordinate with a harmonic bias potential with a force constant of 0.25 kcal/ degree^2 with

windows at 5° spacings. Each window was simulated for 2 ns with a 200 ps equilibration. The PMF was constructed from the umbrella sampling time series using the Weighted Histogram Analysis Method (WHAM).[46–48] In these simulations, temperature was regulated using a Langevin thermostat with a bath temperature of 298.15 K and a damping coefficient of 1 ps^{-1}.

3 Results and Discussion

3.1 Overall Performance

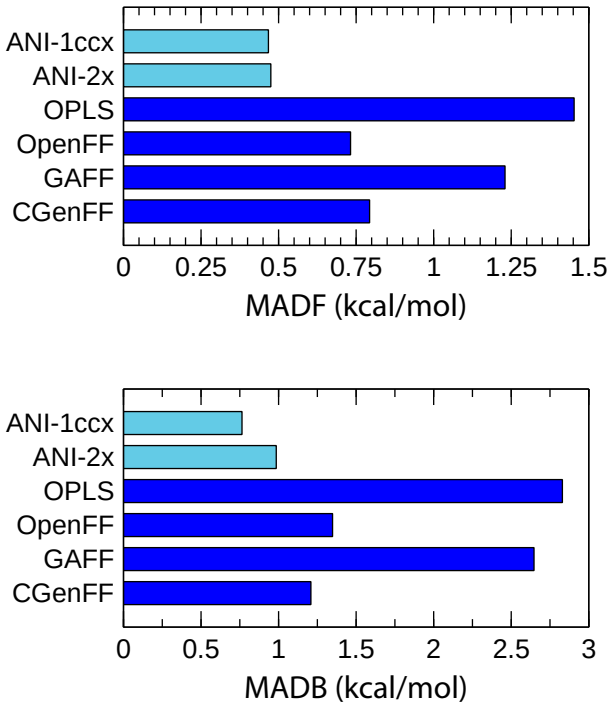

Figure 2: Top: Mean MADF of the PES (Eq. 4) for each method. Bottom: Mean absolute deviation of the rotational barrier height of each method. The CCSD(T1*)/CBS profiles are used as the reference.

Only general trends and some notable examples are discussed here, although the plots of potential energy surfaces of all 88 torsional rotations are included in Supporting Information. We define two metrics for the overall performance of each method. The root-mean-squared-

deviation (MADF) of a method for a given torsion is calculated as the room-mean-square of the potential energy relative to the CCSD(T1)*/CBS//MP2/def2-TZVP reference value, calculated at 5° increments between 0° and 360°,

$$MADF = \frac{1}{n}\sum_{j}^{n}\frac{1}{N_{bins}}\sum_{i}\left(\mathcal{V}_{i,\text{method}} - \mathcal{V}_{i,CCSD}\right)^2 \tag{4}$$

Our second metric is the mean absolute deviation (MADB) of the torsion rotational barrier height for each method, relative to the CCSD(T1)*/CBS//MP2/def2-TZVP barrier height.

$$MADB = \frac{1}{n}\sum_{j}^{n}|\Delta\mathcal{V}^{\ddagger}_{CCSD} - \Delta\mathcal{V}^{\ddagger}| \tag{5}$$

where $\Delta\mathcal{V}^{\ddagger}$ is the barrier height of torsional rotation, defined as the difference between the minimum and maximum energy point on the PES.

The averages of these metrics for the test set are presented in Figure 2. Only profiles that were available for all methods were included in this analysis, so the sulfur-containing compounds (**4**, **6**, **14**, **16**, **24**, **26**, **55**, **56**, and **71**) were not included because the ANI-1ccx NNP is not defined for this element.

For both metrics, the ANI-2x and ANI-1ccx methods outperform all four force fields. Notably, the MADB of ANI-1ccx barrier heights is only 0.7 kcal/mol, indicating that the goal of "sub-kcal" accuracy has already been achieved for these PESs. CGenFF is the most accurate force field, followed by OpenFF, and GAFF. The OPLS force field is the least accurate. The mean signed error (MSE) in the barrier height is positive for all the models except CGenFF, which tends to underestimate barriers by 0.2 kcal/mol. Notably, GAFF has an MSE of 1.2 kcal/mol and OPLS has an MSE of 2.5 kcal/mol, indicating a significant tendency to overestimate torsional barriers.

The relative accuracy of these methods can also be quantified by ranking which method provides the PES with the lowest MADF or the most accurate barrier. These rankings are

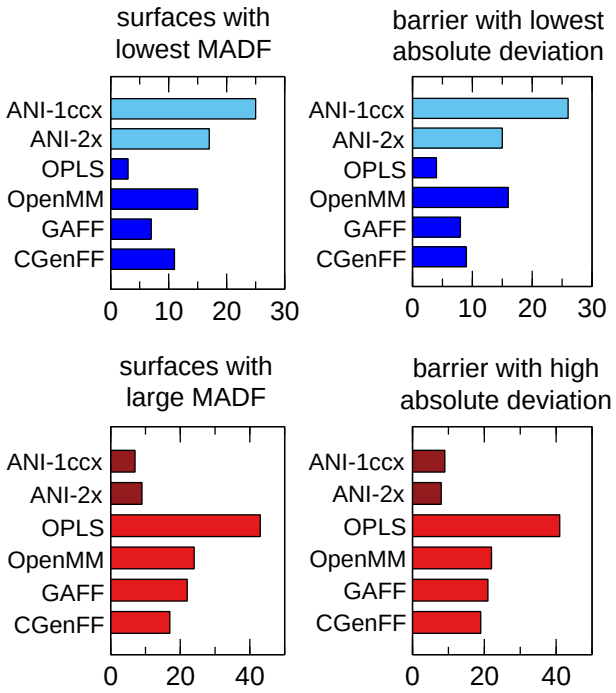

Figure 3: Top left: Number of torsional PESs where a given method has the lowest MADF, Top right: Number of torsional PESs where a given method has the lowest barrier height deviation. Bottom left: Number of torsions where a method gives a high MADF (i.e., on average, more than 1 kcal/mol in error at each point on the PES). Bottom right: Number of torsional PESs where a method predicts the barrier height inaccurately (i.e., more than 2 kcal/mol in error)

presented in Figure 3 (top). By these measures, the ANI-1ccx NNP is most accurate. It should be noted that the ANI-2x method predicts barrier heights with similar results to ANI-1ccx, so it has nearly the same accuracy. although ANI-1ccx is ranked higher because it is generally incrementally more accurate than ANI-2x.

Finally, we can also assess the methods according to the number of torsional PESs where a method performs poorly. These rankings are presented in Figure 3 (bottom). For the MADF criteria, this is defined as a mean absolute deviation (MAD) per point on the surface that is greater than 1 kcal/mol and for the barrier height a poorly performing is defined as one where the predicted barrier height is in error by 2 kcal/mol or more. Based on these metrics for poor performance, the ANI-2x and ANI-1ccx NNP models are also superior compared to the force field models, with the ANI-1ccx methods demonstrating poor performance for only 10 PESs and 10 barriers lower than the best force field (CGenFF). Among the force fields, the CGenFF performs poorly for the fewest torsions, followed by GAFF, OpenFF, and OPLS.

This highlights a major advantage of the NNP methods - the number of cases where they perform poorly is small. The strategy of training these potentials to reproduce molecular energies in general rather than specific interactions results in methods that are robust for PESs outside their training sets. This is a significant advantage for high-throughput screening of protein–ligand binding, where the validation and possible reparameterization of a force field is too time-consuming.

3.2 CGenFF

The CGenFF force field performs better overall than any of the other force fields; however, there are some instances where the barriers are predicted to be much lower than the CCSD(T1)*. This is apparent in N-rich heterocyclics, like **20** and **39**, suggesting that the parameters for C-CA-CA-NA and N-CA-NA-CA dihedrals have a maximum that is too small. For example, the OPLS force field predicts a torsional barrier of **20** more accurately

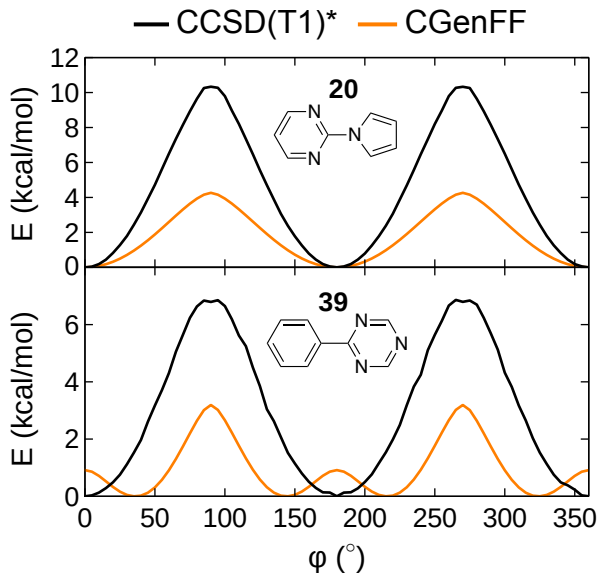

Figure 4: The torsional potential energy surfaces of **20** and **39** are examples where the CGenFF force field is an inaccurate model.

(9 kcal/mol), in part because it uses a C-N-C-N dihedral potential with a maximum of 3.6 kcal/mol instead of the barrier maximum of 1.8 kcal/mol used by CGenFF. Adjustment of a handful of biaryl dihedral terms would improve the accuracy of CGenFF even further.

3.3 OpenFF

OpenFF is the "newest" of the force fields evaluated here and is designed to avoid duplicate or unnecessary parameters. As a result, there are far fewer parameters in the current version of OpenFF compared to the other force fields (i.e., 342 parameters in OpenFF vs more than 6000 parameters for CGenFF). Nevertheless, based on the MADF and barrier height metrics, it generally performs better than GAFF and OPLS for this test set and almost as well as CGenFF. Because relatively few "specific" torsional parameters are currently defined, much of its success is derived from the general biaryl potentials and 1,4 Lennard-Jones and electrostatic parameters. This strategy is less successful for torsions containing aromatic nitrogen atoms in the ortho or ipso positions, such as **10** and **12**. The torsional PESs of these N-containing aromatics are influenced by complex hyper conjugative and electron-

repulsive interactions, which require explicit parameterization for the force field to describe quantitatively.

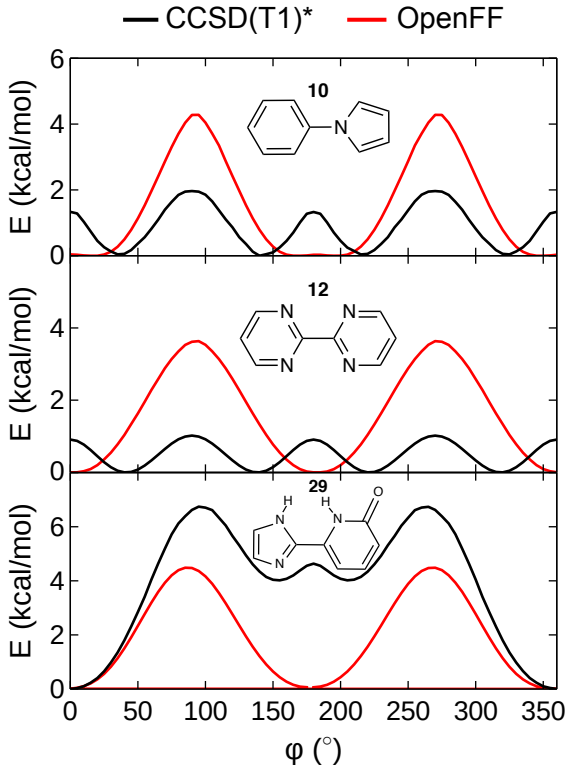

Figure 5: The torsional potential energy surfaces of **10**, **12**, and **29** are examples where the OpenFF force field is an inaccurate model.

3.4 GAFF

An instance where the GAFF force field significantly deviates from the reference PES is where the aryl linkage is through the nitrogen of a pyrrole group. Repulsion between the pyrrole non-bonded pair and the pi system of the benzene ring destabilizes planar conformations, but lone-pair–CH repulsion occurs when the rings are perpendicular, so the minimum energy conformations occur at 30° deviations from planarity. In contrast to this, the GAFF force field predicts a broad minimum corresponding to conformations where the rings are non-planar. The CA–CA–NA–CA torsional parameter is the immediate cause of this issue.

The GAFF force field significantly overestimates the barrier for rotations where there is

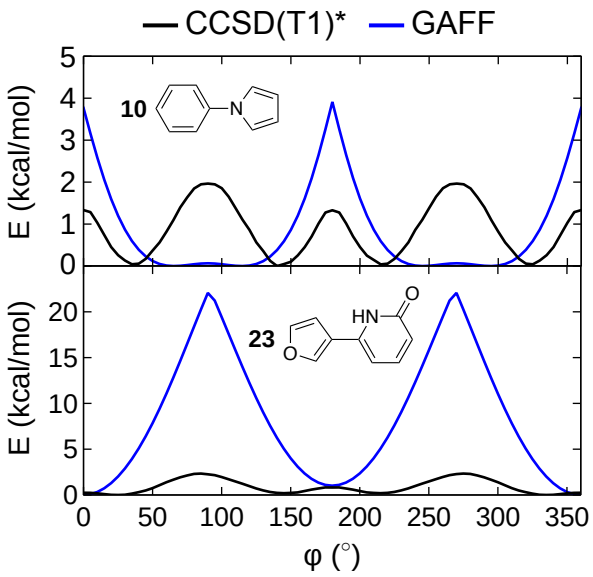

Figure 6: The torsional potential energy surfaces of **10** and **23** are examples where the GAFF force field is an inaccurate model.

an amide NH group in the ortho position of one of the rings. For example, in **23**, GAFF predicts a barrier of 22 kcal/mol, while it is only 2 kcal/mol with CCSD(T1)*. This issue is present in **24**, **25**, **26**, **27**, **28**, and **29**.

3.5 OPLS

The OPLS model performs relatively poorly on this test set. In many cases, this is due to a significant overestimation of the rotational barrier. The mean signed error for the OPLS barrier is 2.5 kcal/mol, indicating that the tendency to overestimate torsional barriers is systematic in this force field. This is evident in the PES for **7** and **12**, where the barrier is overestimated by a factor of **2** and **6**, respectively. In other cases, the topology file generated by the LigParGen server includes torsions that result in asymmetric potential energy surfaces on torsions that should be symmetric. The PES of **9** is an example of this effect.

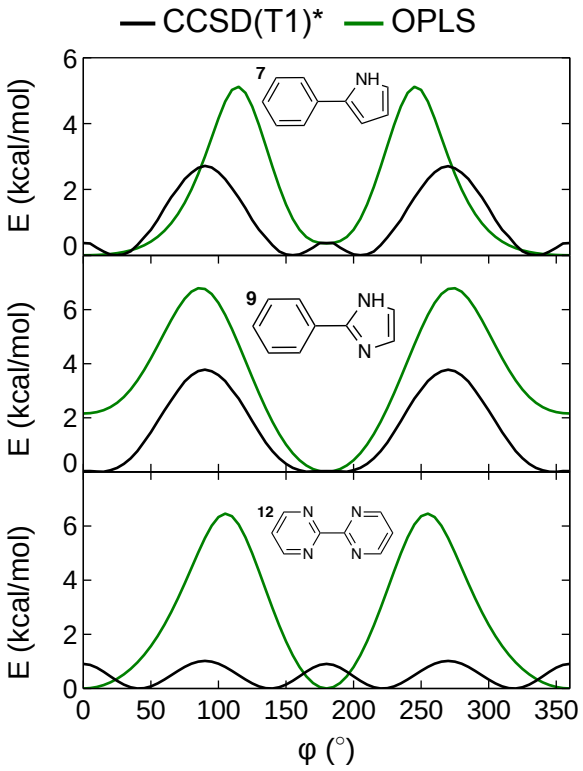

Figure 7: The torsional potential energy surfaces of **7**, **9**, **12** are examples where the OPLS force field is an inaccurate model.

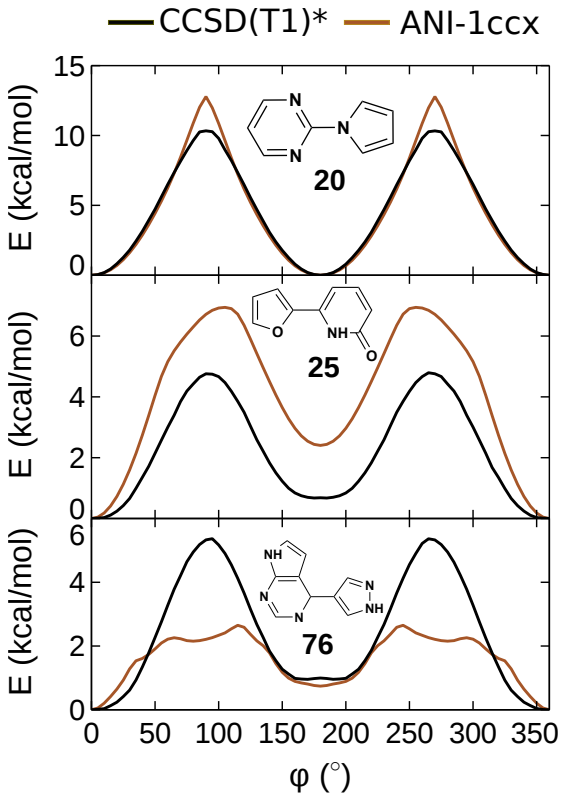

Figure 8: The torsional potential energy surfaces of **20**, **25**, and **76** are examples where the ANI-1ccx force field is an inaccurate model.

3.6 ANI

The ANI-2x and ANI-1ccx NNPs generally outperform the MM models and rarely fail to provide a reasonably accurate surface. The ANI-1ccx NNP gives incrementally greater accuracy than ANI-2x, although it only supports a smaller set of ligands because only the C, N, O, and H elements are defined for it. The relative stability of the cis and trans conformations of **25** is overestimated by the ANI-1ccx potential and the barrier to rotation is significantly overestimated. The PES of **76** is generally irregular and the barrier is significantly underestimated. The lack of explicit electrostatic terms limits the accuracy of these NNPs when the relative stability of two conformations depends on a long-range polar interaction.

3.7 MP2 Failures

Torsional potentials of force fields are often parameterized to reproduce MP2 potential energy surfaces because MP2 is a computationally tractable ab initio method with analytical gradients. Having calculated the energies at both the CCSD(T*)/CBS//MP2/def2-TZVP and MP2/def2-TZVP levels, we can test if these levels of theory provide the same level of accuracy for these torsional profiles. The MP2/def2-TZVP PESs are generally in very good agreement with the CCSD(T1)*/CBS PESs, with a MADF of 0.1 kcal/mol and a MADB of 0.3 kcal/mol.

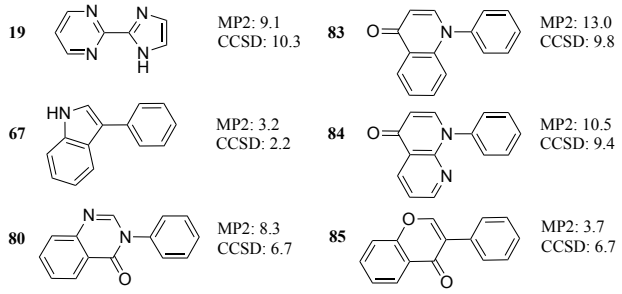

Figure 9: Test set structures where the MP2 barrier height differs by 1 kcal/mol or higher from the CCSD(T1)* surface.

Although the agreement between MP2 and CCSD(T1)* is generally close, of the 88

torsions in the test set, the MP2 and CCSD(T1)* barrier heights differed by more than 1 kcal/mol in 6 instances. In four instances (**80**, **83**, **84**, and **85**), one of the aryls in the rotation was bicyclic, so the transition state occurs when the structure is planar and a steric interaction arises between the ortho hydrogen of the phenyl ring and a non-hydrogen atom in the ortho position of the first ring. The CCSD barrier is lower than the MP2 barrier in 3 out of 4 of these examples, suggesting that MP2 overestimates the strength of this type of steric repulsion. The other examples are nitrogen-containing heteroaromatics (**19**, **67**). The MP2 barrier is lower than the CCSD(T1)* barrier in 1 of these examples, suggesting that there is a small tendency for MP2 to underestimate the electronic repulsion associated with pi-lone pairs. The origin of the deviation in **19** and **67** is less evident, although both are N-containing heteroaromatics.

This raises questions about the common practice of using MP2 potential energy surfaces as the target data for fitting force field torsional potential energy surfaces. Although MP2 is in good agreement with CCSD(T1)* surfaces for most of the molecules, the MP2 and CCSD(T1)* barrier heights differed by 1 kcal/mol or higher for 6 of the torsions. This suggests that parameterizing a force field to reproduce the torsional surfaces of an NNP trained for CCSD(T) data could be advantageous because it is not necessary to calculate CCSD(T1)* potential energy surfaces for each torsion in the ligand.

The ANI-1ccx NNP is incrementally more accurate than the ANI-2x NNP. The ANI-1ccx NNP was developed through a transfer learning approach, where input and output layers of the ANI-1X NNP were retrained using a subset of CCSD*(T1)/CBS data, while the ANI-2x potential was trained to DFT (ωB97x/6-31G*) data exclusively.

3.8 Molecular Dynamics of Torsional Rotations

The large barriers to torsional rotations present in some biaryl compounds can result in large activation energies for conformation isomerization. Consequently, in molecular dynamics simulations of drug compounds containing biaryls, the timescales associated with rotation

around the biaryl bond can be much longer than other, more facile, types of conformational isomerization.

To demonstrate that these methods are practical for use in molecular dynamics simulations for real drug molecules, we have performed simulations of ozanimod in an explicit aqueous solution. Ozanimod is a drug for the treatment of multiple sclerosis that binds the sphingosine 1-phosphate receptor.[49] The rotation around the bond connecting the isopropyl benzonitrile and the diazafuran has a significant barrier height because of the conjugation between the rings and low steric repulsion in planar conformations. The time series of this angle over the course of a 10 ns simulation shows that this molecule undergoes conformational isomerization through rotation around this torsional degree of freedom on the multi-nanosecond timescale.

To investigate this isomerization further, the potential of mean force (PMF) of this degree of freedom was calculated using umbrella sampling and adaptive biasing force[42,43] molecular dynamics simulations. Both methods provide similar PMFs and can be used with the TorchANI NNP/MM interface with NAMD without modification.

Using the umbrella sampling PMF, the rate of isomerization was calculated using Kramers–Smoluchowski transition state theory.[50,51]

$$k_{KS} = D_{TS} \frac{\sqrt{|W''(q_{\text{minimum}}) \cdot W''(q_{TS})|}}{2\pi k_B T} e^{-\Delta W^{\ddagger}/k_B T} \tag{6}$$

In these simulations, the TIP3P-FB water model was used, which predicts a viscosity close to the experimental value,[40] allowing more accurate predictions of diffusion rates in aqueous solutions,[52] and more accurate predictions of the solvent friction on the reaction coordinate.

The diffusion coefficient at the transition state was calculated using the generalized Langevin approach,[53–55] where a strong harmonic potential was used to restrain the simulation to the transition state and the diffusion coefficient was determined by the rate of relaxation of the position autocorrelation function of that time series.

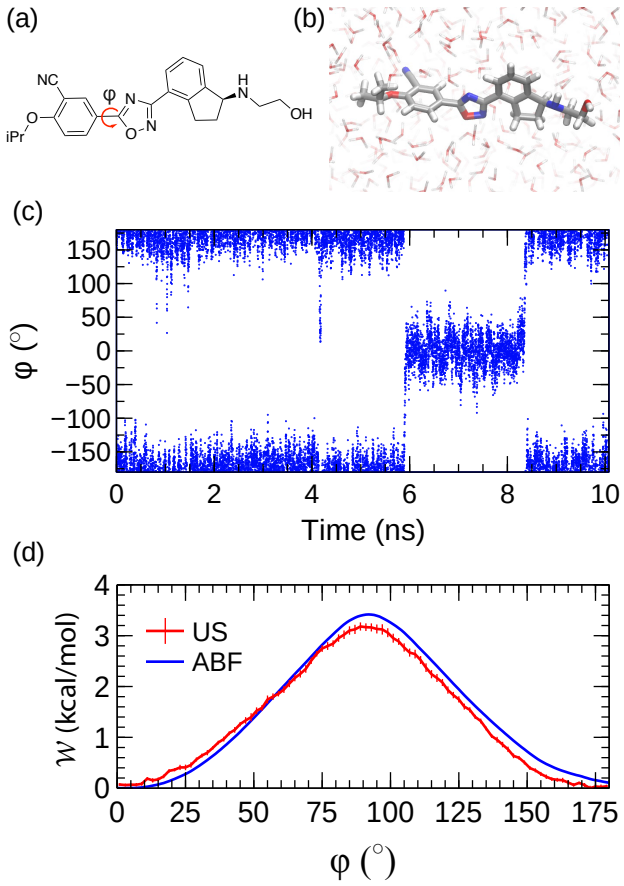

Figure 10: (a) Structure of ozanimod with ϕ torsional angle indicated, (b) Representative structure of ozanimod in solution (c) Time series of ϕ angle from 10 ns MD simulation of ozanimod in aqueous solution (d) ANI-1ccx/TIP3P-FB potential of mean force for rotation around ϕ dihedral.

$$D_{TS} = \frac{\mathrm{var}(q)^2}{\int_0^\infty \langle q(0) \cdot q(\tau) \rangle \mathrm{d}\tau} \tag{7}$$

This theory predicted a rate of isomerization of 4.30×10^{-1} ns^{-1}, which is generally consistent with the slow rates of isomerization observed in the NNP/MM MD simulation (Table 1). These simulations demonstrate that the ANI NNPs can be used immediately to describe the dynamics of slow degrees of freedom of arbitrary drug-like organic molecules using existing simulation methods.

Table 1: Rate theory prediction of isomerization of ozanimod

D_{TS} (rad^2/ns)	67.5
$\Delta W^\ddagger$ (kcal/mol)	3.21
rate (ns^{-1})	4.30×10^{-1}

4 Conclusions

Force field and NNP methods were evaluated for their ability to predict the potential energy surfaces of biaryl torsions found in drug and drug-like molecules ($n = 88$). As these torsions are important features for the structure and dynamics of these molecules, efficient but accurate computational models of these terms are essential for accurate protein–ligand binding simulations. In comparison to high-level ab initio reference data, the ANI-1ccx NNP was the most accurate method and generally predicted barrier heights within 1 kcal/mol. The ANI-2x NNP had a comparable level of accuracy. Significantly, these NNPs provided accurate models in most of the cases and provided poor descriptions in relatively few cases. The robustness and reliability of these NNPs without specific parameterization is particularly useful for simplifying modeling workflows.

The force field methods were less accurate, although there were significant differences in the accuracy of the force fields. The CGenFF was most accurate, followed by the OpenFF, GAFF, and OPLS. The OpenFF model is notable because it performed relatively well despite

having been parameterized with relatively little data and including relatively few parameters. Although MP2 potential energy surfaces were generally in good agreement with the CCSD(T1)* reference values, there were significant differences in 6 instances, suggesting force fields and NNPs should be parameterized to reproduce CCSD(T1)* data for optimal and comprehensive accuracy.

The NNP/MM method was used to simulate the conformational isomerization of the biaryl-containing drug molecule ozanimod. Multi-nanosecond molecular dynamics simulations in an explicit aqueous solvent were performed, as well as umbrella sampling and adaptive biasing force enhanced sampling techniques. These free energy methods can be used in NNP/MM simulations through the NAMD–TorchANI interface, which makes a diverse set of simulation methods available without modification and allows for facile construction. This provides a method for computationally-efficient but highly-accurate models for the intramolecular potential energy surfaces of ligands within biomolecular simulations without relying on a parameterized force field.

Acknowledgement

The authors thank NSERC of Canada for funding through the Discovery Grants program (RGPIN-05795-2016), the School of Graduate Studies at Memorial University for a graduate fellowship, and Dr. Liqin Chen for a scholarship. Computational resources were provided by Compute Canada (RAPI: djk-615-ab). We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

Supporting Information Available

Plots of potential energy surfaces for all methods

References

(1) Wang, J.; Deng, Y.; Roux, B. Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. *Biophys. J.* **2006**, *91*, 2798–2814.

(2) Liebeschuetz, J.; Hennemann, J.; Olsson, T.; Groom, C. R. The good, the bad and the twisted: a survey of ligand geometry in protein crystal structures. *J. Comput. Aided Mol. Des.* **2012**, *26*, 169–183.

(3) Lin, Y.-L.; Meng, Y.; Jiang, W.; Roux, B. Explaining why Gleevec is a specific and potent inhibitor of Abl kinase. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 1664–1669.

(4) Kenno Vanommeslaeghe, O. G.; Mackerell Jr., A. D. Molecular Mechanics. *Curr. Pharm. Des.* **2014**, *20*, 3281–3292.

(5) Gutiérrez Sanfeliciano, S. M.; Schaus, J. M. Rapid assessment of conformational preferences in biaryl and aryl carbonyl fragments. *PLOS ONE* **2018**, *13*, 1–21.

(6) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.

(7) Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **2019**, *15*, 1863–1874.

(8) Mobley, D. L.; Bannan, C. C.; Rizzi, A.; Bayly, C. I.; Chodera, J. D.; Lim, V. T.; Lim, N. M.; Beauchamp, K. A.; Slochower, D. R.; Shirts, M. R.; Gilson, M. K.; Eastman, P. K. Escaping Atom Types in Force Fields Using Direct Chemical Perception. *J. Chem. Theory Comput.* **2018**, *14*, 6076–6092.

(9) Behler, J. Constructing high-dimensional neural network potentials: A tutorial review. *Int. J. Quantum Chem* **2015**, *115*, 1032–1050.

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

(11) Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nat. Comm.* **2019**, *10*, 2903.

(12) Devereux, C.; Smith, J. S.; Davis, 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.* **0**, *0*, null.

(13) Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(14) Lahey, S.-L. J.; Rowley, C. N. Simulating protein–ligand binding with neural network potentials. *Chem. Sci.* **2020**, *11*, 2362–2368.

(15) Rufa, D. A.; Bruce Macdonald, H. E.; Fass, J.; Wieder, M.; Grinaway, P. B.; Roitberg, A. E.; Isayev, O.; Chodera, J. D. Towards chemical accuracy for alchemical free energy calculations with hybrid physics-based machine learning / molecular mechanics potentials. *bioRxiv* **2020**,

(16) Vant, J. W.; Lahey, S.-L. J.; Jana, K.; Shekhar, M.; Sarkar, D.; Munk, B. H.; Kleinekathöfer, U.; Mittal, S.; Rowley, C.; Singharoy, A. Flexible Fitting of Small Molecules into Electron Microscopy Maps Using Molecular Dynamics Simulations with Neural Network Potentials. *J. Chem. Inf. Model.* **2020**, *60*, 2591–2604.

(17) Cole, D.; Mones, L.; Csányi, G. A machine learning based intramolecular potential for a flexible organic molecule. *Faraday Discuss.* **2020**, –.

(18) Dahlgren, M. K.; Schyman, P.; Tirado-Rives, J.; Jorgensen, W. L. Characterization of Biaryl Torsional Energetics and its Treatment in OPLS All-Atom Force Fields. *J. Chem. Inf. Model.* **2013**, *53*, 1191–1199.

(19) Rowley, C. Repository of testset structures and topology files. `https://github.com/RowleyGroup/torsionbenchmark`, 2019.

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

(21) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(22) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(23) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.

(24) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.

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

(26) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(27) Jorgensen, W. L.; Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665–6670.

(28) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **2017**, *45*, W331–W336.

(29) Dodda, L. S.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. *J. Phys. Chem. B* **2017**, *121*, 3864–3870.

(30) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Computational Biology* **2017**, *13*, 1–17.

(31) Gao, X.; Ramezanghorbani, F.; Isayev, O.; Smith, J. S.; Roitberg, A. E. TorchANI: A Free and Open Source PyTorch-Based Deep Learning Implementation of the ANI Neural Network Potentials. *J. Chem. Inf. Model.* **0**, *0*, null.

(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09 Revision D.01. Gaussian Inc. Wallingford CT 2009.

(33) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

(34) Hättig, C. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core–valence and quintuple-ζ basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.

(35) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(36) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci* **2012**, *2*, 73–78.

(37) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F.

Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101.

(38) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; H'enin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 044130.

(39) Rowley, C. NAMD–TorchANI interface scripts. `https://github.com/RowleyGroup/NNP-MM`, 2019.

(40) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.

(41) Koopman, E. A.; Lowe, C. P. Advantages of a Lowe–Andersen thermostat in molecular dynamics simulations. *J. Chem. Phys.* **2006**, *124*, 204103.

(42) Darve, E.; Rodríguez-Gómez, D.; Pohorille, A. Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys.* **2008**, *128*, 144120.

(43) Hénin, J.; Fiorin, G.; Chipot, C.; Klein, M. L. Exploring Multidimensional Free Energy Landscapes Using Time-Dependent Biases on Collective Variables. *J. Chem. Theory Comput.* **2010**, *6*, 35–47.

(44) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.

(45) Johannes, K. Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932–942.

(46) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(47) Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **1995**, *91*, 275–282.

(48) Grossfield, A. WHAM: the weighted histogram analysis method, version 2.0.6, http://membrane.urmc.rochester.edu/content/wham. 2018.

(49) Cohen, J. A.; Arnold, D. L.; Comi, G.; Bar-Or, A.; Gujrathi, S.; Hartung, J. P.; Cravets, M.; Olson, A.; Frohna, P. A.; Selmaj, K. W. Safety and efficacy of the selective sphingosine 1-phosphate receptor modulator ozanimod in relapsing multiple sclerosis (RADIANCE): a randomised, placebo-controlled, phase 2 trial. *Lancet Neurol.* **2016**, *15*, 373 – 381.

(50) Kramers, H. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **1940**, *7*, 284 – 304.

(51) Woolf, T. B.; Roux, B. Conformational Flexibility of o-Phosphorylcholine and o-Phosphorylethanolamine: A Molecular Dynamics Study of Solvation Effects. *J. Am. Chem. Soc.* **1994**, *116*, 5916–5926.

(52) Simulations of lipid bilayers using the CHARMM36 force field with the TIP3P-FB and TIP4P-FB water models. *PeerJ* **2018**, *6*.

(53) Straub, J. E.; Borkovec, M.; Berne, B. J. Calculation of dynamic friction on intramolecular degrees of freedom. *J. Phys. Chem.* **1987**, *91*, 4995–4998.

(54) Hummer, G. Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J. Phys.* **2005**, *7*, 34.

(55) Gaalswyk, K.; Awoonor-Williams, E.; Rowley, C. N. Generalized Langevin Methods for Calculating Transmembrane Diffusivity. *J. Chem. Theory Comput.* **2016**, *12*, 5609–5619.

Graphical TOC Entry