

CLIFF: A component-based, machine-learned, intermolecular force field

Jeffrey B. Schriber,^{1, a)} Daniel R. Nascimento,¹ Alexios Koutsoukas,² Steven A. Spronk,² Daniel L. Cheney,² and C. David Sherrill^{3, b)}

¹⁾*Center for Computational Molecular Science and Technology,
School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia,
30318, USA*

²⁾*Molecular Structure and Design, Bristol Myers Squibb Company, P.O. Box 5400,
Princeton, New Jersey 08543, USA*

³⁾*Center for Computational Molecular Science and Technology,
School of Chemistry and Biochemistry, and School of Computational Science
and Engineering, Georgia Institute of Technology, Atlanta, Georgia, 30318,
USA*

(Dated: 26 April 2021)

Computation of intermolecular interactions is a challenge in drug discovery because accurate *ab initio* techniques are too computationally expensive to be routinely applied to drug-protein models. Classical force fields are more computationally feasible, and force fields designed to match symmetry adapted perturbation theory (SAPT) interaction energies can remain accurate in this context. Unfortunately, the application of such force fields is complicated by the laborious parameterization required for computations on new molecules. Here, we introduce the component-based machine-learned intermolecular force field (CLIFF), which combines accurate, physics-based equations for intermolecular interaction energies with machine-learning models to enable automatic parameterization. CLIFF uses functional forms corresponding to electrostatic, exchange-repulsion, induction/polarization, and London dispersion components in SAPT. Molecule-independent parameters are fit with respect to SAPT2+(3) MP2/aug-cc-pVTZ, and molecule-dependent atomic parameters (atomic widths, atomic multipoles, and Hirshfeld ratios) are obtained from machine learning models developed for C, N, O, H, S, F, Cl, and Br. CLIFF achieves mean absolute errors (MAE) no worse than 0.70 kcal mol⁻¹ in both total and component energies across a diverse dimer test set. For the sidechain-sidechain interaction database derived from protein fragments, CLIFF produces total interaction energies with an MAE of 0.27 kcal mol⁻¹ with respect to reference data, outperforming similar and even more expensive methods. In applications to a set of model drug-protein interactions, CLIFF is able to accurately rank-order ligand binding strengths and achieves less than 10% error with respect to SAPT reference values for most complexes.

^{a)}Electronic mail: jschriber7@gatech.edu

^{b)}Electronic mail: sherrill@gatech.edu (Author to whom correspondence should be addressed)

I. INTRODUCTION

Accurate characterization of noncovalent interactions is required for meaningful predictions of diverse chemical properties, including structures of biomolecules and molecular crystals, protein-ligand interactions, and condensed phase properties. For small-molecule dimers ($\lesssim 40$ atoms), conventional highly-correlated *ab initio* techniques, like coupled-cluster theory, can provide intermolecular interaction energies with benchmark accuracy, and dispersion-corrected density functional theory (DFT-D) remains among the most accurate available supramolecular approaches for much larger systems containing up to hundreds of atoms.^{1–3} Alternatively, symmetry adapted perturbation theory^{4,5} (SAPT) directly computes interaction energies in terms of electrostatic, exchange-repulsion, polarization/induction, and London dispersion components using a hierarchical perturbative scheme. When computationally feasible, SAPT is a particularly attractive candidate for applications in drug binding due to the reasonable accuracy and moderate cost of its most approximate form, SAPT0, and the physical insight provided from component energies.^{6–8} However, for studies that require many energy computations, e.g. dynamics studies or high-throughput screening of drug candidates, quantum mechanical techniques are far too costly, and much cheaper alternatives involving classical or even machine-learned potentials are required. In this work, we present a SAPT-based classical potential designed to produce highly accurate total and component interaction energies for small-molecule and protein-ligand interactions alike.

Characterizing drug binding energetics and relative energies of crystal polymorphs, for example, requires a level of accuracy in intermolecular interaction energies that can be difficult to achieve using standard force fields. One approach to developing force fields with an improved treatment of intermolecular interactions is to use a functional form with energy components that correlate with the components from SAPT.^{9–12} These SAPT-based force fields provide a good framework for computing accurate interaction energies, and their accuracy has been demonstrated in myriad dynamics studies on diverse molecules.^{10,11,13–16} In practice, functional forms in SAPT-based force fields typically require numerous monomer-specific atomic parameters, including multipole moments and polarizabilities, for each atom in each monomer of interest, in addition to a set of interaction parameters that arise in the functional forms and are determined by fitting to total or component energies. Throughout this work, we refer to this second class of interaction parameters as *global* parameters since they are inherent to the force field and do not change when applied to different molecules. The atomic parameters can be computed from *ab initio* monomer densities

in tandem with a charge partitioning scheme.^{11,17–19} While most SAPT-based force fields have sets of potentially numerous global parameters, replacing these parameters that require fitting with ones derived from *ab initio* densities where possible can greatly increase a model’s transferability and robustness.²⁰ Atomic multipoles or point charges are commonly obtained in this way by computation of a molecular density followed by a distributed multipole analysis (DMA).^{21,22} Recently developed force fields use more information from molecular densities to parameterize advanced functional forms.^{18,19,23–25} For example, MASTIFF uses approximated atomic density decay rates computed from *ab initio* molecular densities to compute atomic density overlaps useful for treating short-range effects.^{26,27} MEDFF also extensively uses atomic parameters derived from *ab initio* densities, and as a result only requires three global fitting parameters in the entire model.²⁸

While computation of molecular densities for small-molecule parameterization is straightforward, the computation of densities in molecules containing hundreds to thousands of atoms is nontrivial and at times prohibitive. An alternative approach, used in the AMOEBA force field, is to use a hybrid model, where *ab initio* densities are used when feasible and tabulated, atom-typed parameters derived from fragment *ab initio* computations are used for large molecules, proteins, and nucleic acids.^{29–33} This approach has been very successful for a variety of biomolecular applications, even including transition metals.^{34,35} However, the requirement either (1) to recompute an *ab initio* molecular density or (2) to refit numerous atomic parameters in computing any new type of molecule poses a significant challenge for applications requiring large numbers of computations on different molecules, for example the computational screening of large (100 atom) drug targets in a protein binding region.

To achieve transferability and to avoid numerous reparameterizations, recently proposed alternatives to high-accuracy force fields use machine learning (ML) either in combination with or in replacement of physics-based functional forms.^{36–39} “Pure” ML models use only molecular coordinates as input to predict, most commonly, total energies,^{40–49} and recent examples from our group can instead target interaction energies very accurately.^{50,51} The benefit of these approaches is that transferability of the model is limited only by the extent of training data used to build the model, and no parameterization or quantum-mechanically derived electronic densities are required for computation of a new molecular system. While curation of training data is a significant challenge in designing a maximally-transferable model, the demonstrated accuracy and performance of modern ML potentials makes them a very attractive option.

On the other hand, at long range at least, there are well-known functional forms that accurately

describe intermolecular interactions, and it seems profitable to take advantage of them. Thus, an alternative solution to transferability is to use physics-based models where some or all of the parameterization is done using ML.^{52–54} One manifestation of this approach is the IPML method introduced by Bereau et al.,⁵⁵ where ML models are used to compute required atomic parameters for functional forms used in existing force fields. An attractive consequence of using ML for only the atomic parameters is that the ML models need to be only trained on monomers, thus avoiding complications in generating very large databases of dimer interaction energies, assuming any global interaction parameters can be obtained by fitting with a relatively small dimer database. With only eight global fitting parameters in the functional forms, IPML is able to predict intermolecular interaction energies with high accuracy for diverse small-molecule neutral dimers composed of C, N, O, and H using a single parameter set. Without the requirement of computing electron densities for new computations, IPML can in principle be very easily applied to larger molecules. Despite having terms that correspond to SAPT components, the internal parameters in IPML are fit with respect to total energies only, rendering the individual components inaccurate and unable to provide qualitative insight. Furthermore, inaccurate component energies mean that accurate total interaction energies rely on a systematic cancellation of errors, which is not guaranteed across diverse interaction types. For example, when applied to the supramolecular complexes in the S12L database, IPML only shows reasonable agreement with reference total interaction energies when these dimers are included in fitting the global interaction parameters.⁵⁵

We were inspired by IPML’s interesting combination of a sound, physics-based force field with ML to quickly and automatically determine most required atomic parameters. Here, we introduce a new intermolecular potential called the component-based, machine-learned intermolecular force field (CLIFF), which uses ML models to provide atomic parameters for a SAPT-based force field. The primary goal of CLIFF is to compute both component energies as well as total energies for non-covalently bound systems to an accuracy comparable to that of high-level *ab initio* theories. With accurate components, CLIFF relies minimally on cancellation of errors among component energies, and can provide useful component-based analysis valuable in SAPT applications.⁷ Despite extensive use of machine-learned atomic parameters, our functional forms still contain global parameters that require fitting. Each component equation in CLIFF contains global parameters unique to that component, which allows them to be optimized with respect to SAPT component energies in addition to total energies. Furthermore, this parameterization is only done once and not during the computation on new molecular systems, and we will show it does not significantly

limit our transferability. By using ML for all parameterizations of new molecules, we are able to compute intermolecular interaction energies for diverse small-molecule benchmark dimers and large protein-ligand complexes with the accuracy of high-level SAPT-based force fields. We additionally improve upon IPML’s ability to model organic molecules containing C, N, O, and H atoms by adding parameterizations and ML models for molecules containing S, F, Cl, and Br, to enable more extensive applications in biochemistry and drug discovery.

II. THEORY

A. Atomic Parameterization from Machine Learning

Similar to the approaches of Van Vleet et al.^{26,27} and Vandenbrande et al.,²⁸ CLIFF uses component equations that make use of electronic density overlaps. To write component equations in a pairwise-atomic form, a partitioning of the monomer electron densities to their atoms is required and is achieved by using atoms-in-molecules (AIM) methods.^{9,56,57} Resultant AIM densities are atom-centered electronic distributions that are inherently aware of their local chemical environments. Force field parameters derived from AIM densities are consequently specific to the molecules under study, without requiring any refitting of large numbers of global parameters, and provide a means to develop accurate and transferable models. Atomic multipoles, atomic widths, and Hirshfeld ratios are the atomic properties used in CLIFF, and they can all be derived from AIM densities.

To partition the molecular electronic density, we use the minimal basis iterative stockholder (MBIS) method.⁵⁸ MBIS represents the molecular electronic density as a sum of individual atomic densities defined using sums of atom-centered Slater functions, where the sums run over the shells of the particular atom. From this density, atomic multipole coefficients can be computed, and the Slater widths of the outer shell define the valence width, referred to in this work as the atomic width. Multipole coefficients from MBIS are used in the electrostatic model, as well as in computing the induced dipoles for the induction term. The atomic widths approximate the exponential decay rate of the electronic density and are used to compute an approximation to the overlap of atomic electron densities. As discussed later, these overlap terms are used throughout CLIFF, particularly to model short-ranged interactions.

In addition to MBIS partitioning, we also use a Hirshfeld partitioning to compute effective

volume ratios, h_i , defined as the ratio of the effective AIM volume to the effective volume of the free atom,

$$h_i = \frac{V_i^{\text{AIM}}}{V_i^{\text{free}}} \quad (1)$$

We denote these volume ratios as Hirshfeld ratios due to the partitioning scheme used.⁵⁹ Similar quantities can be computed with MBIS densities, though we chose the Hirshfeld partitioning for consistency with the Tkatchenko-Scheffler⁶⁰ method used later, though evidence suggests that MBIS-derived quantities work equally well in this context.^{28,58}

To compute the atomic multipoles, atomic widths, and Hirshfeld ratios, CLIFF uses separate ML models for each chemical element used (C, H, N, O, S, F, Cl, Br), for a total of 24 models. All models use kernel-ridge regression (KRR),^{61,62} which was initially chosen for its simplicity. KRR predicts a desired quantity, $\mathbf{y}$, by applying a kernel function, $k(\mathbf{x} \cdot \mathbf{x})$ to a set of regression coefficients, α , determined using training data,

$$\mathbf{y} = \sum_i^N k(\mathbf{x}_i \cdot \mathbf{x}) \alpha_i \quad (2)$$

where the sum runs over N training points, and $\mathbf{x}$ refers to the discretized representation of the atom on which we are predicting, and $\mathbf{x}_i$ is similarly the descriptor for an atom used in training. The kernel function in this work computes the similarity between a pair of descriptors using the usual Laplacian kernel involving one hyperparameter, γ . The regression coefficients are determined analytically from the training data, $\alpha = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y}$, where $\mathbf{K}$ is the kernel matrix whose elements are the kernel function applied to all pairs of training data points, $\mathbf{y}$ contains the reference property values, and λ is a regularization hyperparameter that requires fitting. IPML uses Coulomb matrices⁶¹ as descriptors for models that compute isotropic quantities (atomic widths and Hirshfeld ratios), and uses the atomic spectrum of London and Axilrod-Teller-Muto potentials^{62,63} (aSLATM) as the representation for computing multipoles.⁵⁵ Coulomb matrices are significantly more efficient than aSLATM but are unable to easily predict orientation-dependent quantities like atomic dipoles and quadrupoles. In our experience, Coulomb matrices were unable to give reliable predictions for even isotropic quantities when building models for heavier elements. Therefore, we use the aSLATM representation implemented in QML⁶⁴ in computing atomic multipoles, atomic widths, and Hirshfeld ratios for all chemical elements used.

While CLIFF is designed to compute intermolecular interaction energies, the ML models for atomic properties do not need to be trained on dimers, as they all depend only on monomer densities. Our monomer database is composed entirely of structures obtained from the ChEMBL

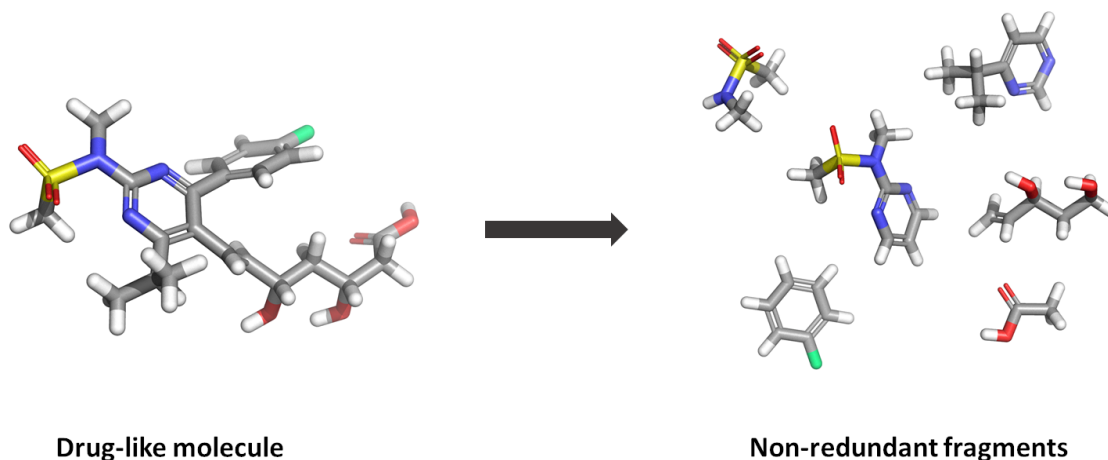


FIG. 1. Depiction of the fragmentation scheme used to generate the monomer database for all ML models. Large drug-like molecules are fragmented to a set of non-redundant subunits.

database,⁶⁵ which we filtered to a subset of approximately 872,000 drug-like molecules. Since the atomic properties we compute are local in nature, we chose to fragment this set into molecules of 5 to 12 heavy atoms, as depicted in Figure 1. Unique fragments are curated to obtain a set of 8,138 chemically diverse molecules with moieties representative of drug-like molecules and protein targets. We compute all reference densities with PBE0/aug-cc-pV(D+d)Z using PSI4.⁶⁶ Reference atomic properties are computed using MBIS and Hirshfeld routines implemented in HORTON.⁶⁷

In building ML models for multipoles, atomic widths, and Hirshfeld ratios, randomized subsets of the 8,138 molecules are taken for each element, with a different percentage of the database being used for each element due to a higher representation of certain elements (C,H) compared to others (S, Cl, F, Br). Construction of the ML models is based on the IPML implementation, and we used our training datasets to re-fit KRR hyperparameters. With the ML models trained, we predict a single set of atomic properties for each monomer to be used in all functional forms. We first generate the aSLATM descriptors to encode local environments in our atomic representations, and we then use these descriptors in all KRR models to predict the desired properties. Identically to IPML, we scale the predicted atomic charges so that the net charge of each monomer is exactly zero. This scaling avoids spurious electrostatic interactions between monomers. As will be shown, our atomic charges are sufficiently accurate so that we do not expect any qualitative difference between our predicted charges and those from MBIS.

B. Component Functional Forms

1. Electrostatics

An accurate model of electrostatics needs the flexibility to describe the different physical interactions that occur at both long and short range. Long-range interactions are well described by a distributed multipole expansion, in which atomic multipole coefficients, M_i for atom i in molecule A and M_j for atom j in molecule B , are contracted with the point multipole interaction tensor, T_{ij} , to produce the electrostatic energy,⁶⁸

$$E_{\text{elst}} = \sum_{i \in A} \sum_{j \in B} M_i^T T_{ij} M_j \quad (3)$$

The multipole expansion alone, however, fails at short range (~ 2 times the van der Waals radius) due to charge penetration effects defined as the enhanced electrostatic interaction resultant from overlapping electron densities and reduced nuclear screening.⁶⁸ Charge penetration is missing from the multipole expansion because it uses point multipoles rather than multipoles with some representation of widths required to model overlapping charge densities. Correspondingly, charge penetration corrections typically employ damping functions to give atomic charges widths necessary to model reduced nuclear screening.

In this work, we employ a damped multipole electrostatic model recently introduced by Rackers et al.⁶⁹ While most charge penetration corrections only include damping functions for atomic charges, Rackers' model uses a damped interaction tensor to include charge penetration effects from all atomic multipoles. We can write the electrostatic energy by separating the nuclear charge (Z_i for atom i) from the corresponding multipoles,

$$E_{\text{elst}} = \sum_{i \in A} \sum_{j \in B} \frac{Z_i Z_j}{r_{ij}} + Z_i T_{ij}^{f_1} M_j + M_i^T T_{ij}^{f_1} Z_j + M_i^T T_{ij}^{f_2} M_j \quad (4)$$

where $T_{ij}^{f_1}$ and $T_{ij}^{f_2}$ are the damped interaction tensors for nuclear-multipole and multipole-multipole interactions, respectively. We adopt the same damped interaction tensors as Ref 69, where the nuclear-multipole damping function is defined as

$$f_1(r_{ij}) = 1 - e^{-K_i^{\text{elst}} r_{ij}} \quad (5)$$

and the multipole-multipole damping function is,

$$f_2(r_{ij}) = 1 - \frac{(K_i^{\text{elst}})^2}{(K_i^{\text{elst}})^2 + (K_j^{\text{elst}})^2} e^{-K_i^{\text{elst}} r_{ij}} - \frac{(K_j^{\text{elst}})^2}{(K_j^{\text{elst}})^2 + (K_i^{\text{elst}})^2} e^{-K_j^{\text{elst}} r_{ij}} \quad (6)$$

Elements of the damped interaction tensors are computed by taking successive partial derivatives of $f_1(r_{ij})r_{ij}^{-1}$ for $T_{ij}^{f_1}$ and $f_2(r_{ij})r_{ij}^{-1}$ for $T_{ij}^{f_2}$, analogous to the undamped tensor where these derivatives are applied only to r_{ij}^{-1} . In the damping functions, we introduce atomic fitting parameters K_i^{elst} which are obtained by fitting equation (4) to SAPT dimer computations. Rather than relying on additional fitting or quantum chemical calculations, our multipoles up to quadrupoles are obtained from the previously mentioned machine learning models. As discussed in more detail later, the K_i^{elst} are atom typed largely by element and number of bonding partners.

2. Exchange-repulsion

Exchange-repulsion, frequently referred to as Pauli repulsion, is a repulsive force resulting from monomer wave function overlap under the requirement that the total dimer wave function be antisymmetric. Consequently, many classical models of exchange-repulsion make use of some approximation to the overlap of electronic densities of each monomer.^{70–72} In fact, empirical findings suggest that exchange-repulsion depends linearly on the density overlap, S_{AB} ,

$$E^{\text{exch}} = KS_{AB} \quad (7)$$

for monomers A and B with K being a yet-undefined proportionality constant.⁷⁰ In the context of advanced force fields,^{26,58,73–76} the overlap is defined using an atoms-in-molecules decomposition,

$$E^{\text{exch}} = \sum_{i \in A} \sum_{j \in B} K_{ij} S_{ij} \quad (8)$$

where S_{ij} is the overlap in electronic density between atom i on monomer A and atom j on monomer B . Following previous work,²⁶ we define the proportionality constant with a multiplicative combining rule, $K_{ij}^{\text{exch}} = K_i^{\text{exch}} K_j^{\text{exch}}$ for atom-typed parameters K_i^{exch} .

Numerous definitions of S_{ij} have been proposed, and they typically require some definition of an atomic charge width or charge decay rate. The IPML model uses the expression from Vandenberg et al.,²⁸

$$S_{ij} = \frac{N_i N_j}{8 r_{ij}} \left[\left(\frac{4}{\left(\frac{r_i^2}{2} + \frac{r_j^2}{2} \right)^3} + \frac{r_{ij}}{\left(\frac{r_i^2}{2} + \frac{r_j^2}{2} \right)^2} \right) e^{-\frac{r_{ij}}{r_i}} + \left(\frac{4}{\left(\frac{r_i^2}{2} + \frac{r_j^2}{2} \right)^3} + \frac{r_{ij}}{\left(\frac{r_i^2}{2} + \frac{r_j^2}{2} \right)^2} \right) e^{-\frac{r_{ij}}{r_j}} \right] \quad (9)$$

which uses valence atomic populations, N_i , and valence widths, r_i , both derived from an MBIS procedure. Unfortunately, the valence widths of two atoms, particularly of the same element, are

oftentimes very similar and cause equation (9) to become unstable, requiring separate approximations to handle these cases. We opt for a simpler overlap approximation developed by Van Vleet et al.^{26,27} wherein an effective width,

$$B_{ij} = \frac{1}{i+j} \quad (10)$$

is used in the expression for S_{ij} extrapolated to the $i = j$ limit,

$$S_{ij} = [\frac{1}{3}(B_{ij}r_{ij})^2 + B_{ij}r_{ij} + 1]e^{-B_{ij}r_{ij}} \quad (11)$$

Equation (11) is more numerically stable and computationally efficient than equation (9), and it has been shown to incur nearly negligible errors with respect to the exact overlap expression.²⁶ Our model for exchange is thus defined as equation (8) using the atomic overlap definition from equation (11).

3. Dispersion

Coupled dynamical correlations between electrons on different monomers can result in an attractive interaction, defined as dispersion. While dispersion interactions can be treated with *ab initio* techniques, empirical models for dispersion have gained popularity for use with classical potentials, DFT, and wave function-based techniques alike due to their significantly reduced cost and acceptable accuracy.^{2,60,77–85} Many classical dispersion models rely on the well-known dispersion series,

$$E_{\text{disp}} = -\sum_{i \in A} \sum_{j \in B} \sum_{n=6}^{\infty} \frac{C_n(i,j)}{r^{2n}} \quad (12)$$

where dispersion coefficients $C_n(i,j)$ are typically computed from imaginary-frequency-dependent polarizabilities and $i(j)$ is an atom in monomer $A(B)$. This series is derived from a second-order perturbative analysis of the interaction energy between dimers with zero charge overlap, and thus requires an additional short-range treatment usually in the form of damping functions.⁸⁶

In this work, we use a Tang-Toennies damped dispersion model⁸⁶ with dispersion coefficients derived from machine-learned atomic properties. Following the Tkatchenko-Scheffler (TS) method,⁶⁰ the atomic-pairwise $C_6(i,j)$ coefficients are computed using single-atom $C_6(i)$ coefficients and atomic polarizabilities, α_i ,

$$C_6(i,j) = \frac{2C_6(i)C_6(j)}{\alpha_i C_6(i) + \alpha_j C_6(j)} \quad (13)$$

We define the single-atom dispersion coefficients using free-atom coefficients⁸⁷ and the machine-learned Hirshfeld ratios, $C_{6i} = C_{6i}^{\text{free}} h_i^2$. Similarly, atomic polarizabilities are computed from free-atom polarizabilities and Hirshfeld ratios, $\alpha_i = \alpha_i^{\text{free}} h_i$, where both C_{6i}^{free} and α_i^{free} are computed from free-atom densities and tabulated for each element type. To compute higher-order dispersion coefficients, we rely on recursion relations found in the D corrections introduced by Grimme and coworkers.^{2,79,80} The C_{8ij} coefficients require the C_{6ij} coefficients,

$$C_{8ij} = 3C_{6ij} \sqrt{Q_i Q_j} \quad (14)$$

where $Q_i = \overline{Z}_i \frac{r_i^4}{r_i^2}$ and r_i^n is a simple multipole-like expectation value that is computed from free-atom densities and also tabulated for each element. The C_{10ij} dispersion coefficients follow a much simpler relation,

$$C_{10ij} = \frac{49}{40} \frac{C_{8ij}^2}{C_{6ij}} \quad (15)$$

To avoid divergences as r approaches zero, damping functions for each term in the dispersion series are required. We use the Tang-Toennies damping function,

$$f_n = 1 - \left(\sum_{k=0}^n \frac{x_{ij}^k}{k!} \right) e^{-x_{ij}} \quad (16)$$

where the argument, x_{ij} , is related to the overlap of atomic densities (S_{ij}) and can be written in terms of the effective widths, B_{ij} , defined in equation 10,

$$x_{ij} = B_{ij} r_{ij} + \frac{2B_{ij}^2 + 3B_{ij}}{(B_{ij} r_{ij})^2 + 3B_{ij} r_{ij} + 3} r_{ij} \quad (17)$$

as initially derived by Van Vleet et al.²⁶ An extremely appealing result in this procedure is that no fitting is required in the damping function itself, so long as both the machine-learned widths and the reference widths computed from MBIS for training are accurate.

Computation of C_{6ij} coefficients using the Tkatchenko-Scheffler method with AIM-derived quantities was initially reported by Verstraelen et al.,⁵⁸ and the use of this methodology in computing dispersion corrections was done by Vandebrande et al.²⁸ in the MEDFF method. In the spirit of IPML, our approach follows this work by machine learning the required atomic parameters rather than computing quantum mechanically derived densities and performing a subsequent AIM routine. Our approach also differs slightly from MEDFF in the functional forms, where CLIFF uses different definitions of damping functions and higher-order dispersion coefficients. Lastly, to compensate for slight overestimation of the higher-order dispersion coefficients, we fit

their contribution with atomic pairwise parameters, $K_{ij}^{\text{disp}} = K_i^{\text{disp}} K_j^{\text{disp}}$, when computing the total dispersion,

$$E_{\text{disp}} = \sum_{i \in A} \sum_{j \in B} \left(\frac{C_{6ij}}{r_{ij}^6} f_6(r_{ij}) + K_{ij}^{\text{disp}} \sum_{n=8}^{10} \frac{C_{nij}}{r_{ij}^n} f_n(r_{ij}) \right) \quad (18)$$

Scaling terms for which $n \geq 6$ is also done in D corrections developed by Grimme et al.,⁷⁹ though we use atom-typed parameters to account for atom-dependent errors in higher-order dispersion coefficients.

4. Induction

To compute the induction component, also referred to as the polarization energy, we combine the commonly used Thole method⁸⁸ with a short-range correction defined using atomic density overlaps,

$$E_{\text{ind}} = \sum_{i \in A} \sum_{j \in B} \mu_i T_{ij} M_j + K_{ij}^{\text{indu}} S_{ij} \quad (19)$$

where μ_i are induced atomic dipoles, S_{ij} is the atomic density overlap defined from equation (11), and $K_{ij}^{\text{indu}} = K_i^{\text{indu}} K_j^{\text{indu}}$ are fitting parameters. Following the procedure of Ren and Ponder,^{29,89} the induced dipoles for atom i are computed in an iterative, self-consistent procedure defined by,

$$\mu_i(n+1) = (1 - \alpha_i) \mu_i(n) + \alpha_i \left[\mu_i(0) + \sum_{\substack{k \in A \cup B \\ k \neq i}} T_{ik} M_k \right] \quad (20)$$

where k labels all atomic sites on all monomers other than i . We use a successive over-relaxation procedure to more quickly converge equation (20) by setting $\alpha_i = 0.7$. The “direct” induced dipole is defined as

$$\mu_i(0) = \sum_{j \in B} T_{ij} M_j \quad (21)$$

where the summation is only between sites on different monomers. The polarizabilities (α_i) are computed as discussed previously, and come from machine-learned Hirshfeld ratios and tabulated free-atom polarizabilities. The interaction tensor, T_{ij} uses Thole damping to smear atomic charge distributions,

$$f_{\text{Thole}} = \frac{3a}{4} e^{-au^3} \quad (22)$$

where $u = r_{ij} (\alpha_i \alpha_j)^{\frac{1}{6}}$, and a is a unitless smearing coefficient conventionally defined as 0.39. We instead re-fit the smearing coefficient together with our K^{indu} parameters. An interesting outcome in our approach, and that of IPML, is that there is no fitting required to compute either the

multipoles or the polarizabilities in the first (polarization) term in equation (19). Our approach of combining the usual polarization term with a short-range, overlap-based correction was previously introduced in the MASTIFF force field with good results.²⁶ Correcting a polarization energy with an overlap-based term is also motivated by an apparent relationship between first-order exchange energies and charge-transfer interactions that manifest in short-range induction interactions.^{9,10,90}

C. Global Interaction Parameters

As demonstrated, parameters used by CLIFF come in two categories. First are the set of monomer-specific *atomic* parameters—atomic multipoles, atomic widths, and Hirshfeld ratios—computed using ML models. Second are what we refer to as *global* interaction parameters, K_i , which are determined by fitting to dimer interaction energies and remain constant in applying CLIFF to new dimers. Each interaction energy component has a set of seventeen atom-typed global parameters, K_i , and the smearing coefficient, a , in induction is also fit but is not itself atom-typed. In defining atom types, we find a balance between accuracy and fitting stability by defining a modest number of atom types using the atomic number and coordination number of a given atom. One exception is with hydrogen, where we define a separate atom type based on the element of its bonding partner. All atom types and global interaction parameters are summarized in Table I.

Following many SAPT-based force fields, we can fit the global interaction parameters using their corresponding SAPT component energy, with or without a total interaction energy. An advantage of combining ML atomic parameterization with physics-based functional forms is that the number of reference interaction energies does not need to be nearly as extensive in comparison to pure ML approaches. As a result, high-level reference data is feasible for use in fitting global parameters as the number of required dimers can remain comparatively small. Based on its performance in comparison to CCSD(T)/CBS total interaction energies,⁹¹ we fit our interaction parameters to the component and total energies from SAPT2+(3) MP2/aug-cc-pVTZ reference calculations. One significant complication in this choice is that the distinction between electrostatics, exchange, induction, and dispersion becomes less clear due to numerous cross-terms, though we adopt the canonical partitioning of these terms summarized in ref 91. Specifically, this partitioning categorizes the HF and MP2 terms within the induction term.

Reference interaction energies used for fitting were computed with dimers from a custom-built database. Our dimer database is composed of two collections of dimers. For the first collection, we

TABLE I. Atom-typed global interaction parameters used in CLIFF. These parameters are fit with respect to SAPT2+(3) MP2/aug-cc-pVTZ reference energies using $\epsilon = 0.4$ in our fitting function. Along with these atom-typed parameters, fitting of the smearing coefficient from equation 22 resulted in a value of 0.38539.

Atom Type	K^{elst}	K^{exch}	K^{indu}	K^{disp}
C4	3.3911	2.2649	$2.1196 \cdot 10^{-5}$	0.3489
C3	3.3323	2.4566	0.2841	0.3801
C2	3.1057	2.8023	0.7843	0.4747
N3	3.4371	4.4660	1.7546	0.2515
N2	3.0371	4.6251	1.5211	0.9213
N1	3.3785	3.4896	0.8137	0.8142
O2	3.8700	5.8538	1.1477	0.5480
O1	3.6031	5.3435	1.6372	0.7794
S2	3.1034	3.2842	0.7699	0.6898
S1	3.0618	3.1773	0.9862	0.7250
HC	3.5982	0.9890	0.3781	0.1619
HN	3.2554	0.6910	0.5952	0.1420
HO	3.1255	0.5996	0.6856	$8.0387 \cdot 10^{-7}$
HS	3.5974	0.7909	0.6031	$4.2571 \cdot 10^{-6}$
F	4.3157	7.6036	1.5281	0.5935
Cl	3.4402	3.8152	0.8468	0.6289
Br	3.6942	4.1008	1.1612	0.4993

take all possible pairs of thirty interacting sites on twenty-three unique monomers (Fig. 2), generating an initial set of 465 dimers. Dimer configurations are derived by pairing the monomers to interact through specific sites, labeled as red atoms and dots in Fig. 2. Angles and dihedral angles involving the interaction sites and their neighboring atoms are sampled randomly, as described in more detail in Ref 50. After determining the rotational orientation, a range of intermolecular distances is generated by varying the van der Waals overlap, defined using the van der Waals radii of the interacting sites. Specifically, the van der Waals overlap is varied from $1.0 \text{ \AA}$ to $-1.0 \text{ \AA}$ in $0.1 \text{ \AA}$ steps, where positive values indicate separation between the monomers and negative values indicate clashes. For dimers formed from strong electrostatic interactions, including hydrogen or

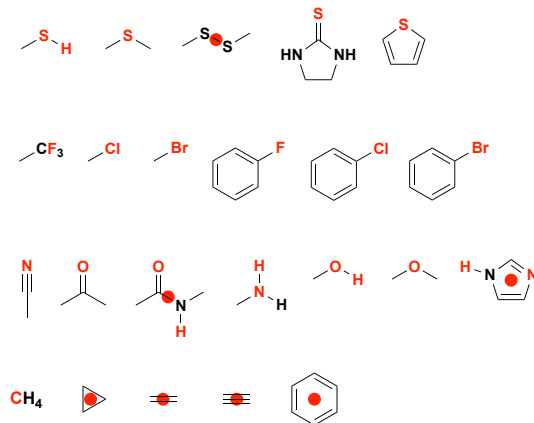


FIG. 2. Monomers used to construct one collection of the dimer database. Interaction sites can be atoms in the molecule, labeled in red, or they can be delocalized over a number of bonds, labeled with a red dot.

halogen bonds, we lower the minimum distance to $1.3 \text{ \AA}$. For each distance, three configurations with different angular orientations are obtained to yield 63-72 configurations for each dimer. This procedure ensures that our training set includes dimers with relatively short intermolecular distances. High representation of short-range interactions is important in fitting, as the global interaction parameters are generally linked to short-range corrections in CLIFF. From this set of roughly 30,000 structures, we take a subset of 7,000 structures chosen from an even distribution of total interaction energies.

While the first dimer collection emphasized thorough configurational sampling of dimers containing all atom-types in CLIFF, the second collection ensures ample representation of equilibrium, and at times strongly-attractive interactions. Dimers in this collection were selected from a large database representing interactions between drug molecules and proteins, and they were constructed using a previously described procedure ultimately resulting in 884 dimers used in training.⁵⁰ All dimers are at or near equilibrium geometries, and they span a wide range of energies (-0.3 to $-27.5 \text{ kcal mol}^{-1}$) and chemical structures. Details of this database are presented in the SI. For the complete training set used in CLIFF, we combine all dimers from this collection with the 7,000 dimers from the first collection resulting in 7,884 dimers in total.

Rather than fitting purely to total or component energies, we use a multi-target fitting function designed to minimize mean squared errors (MSE) of both total and component energies,

$$\mathcal{L} = (1 - \alpha) \text{MSE}(E_{\text{total}}) + \alpha \sum_{i \in C} \text{MSE}(E_i) \quad (23)$$

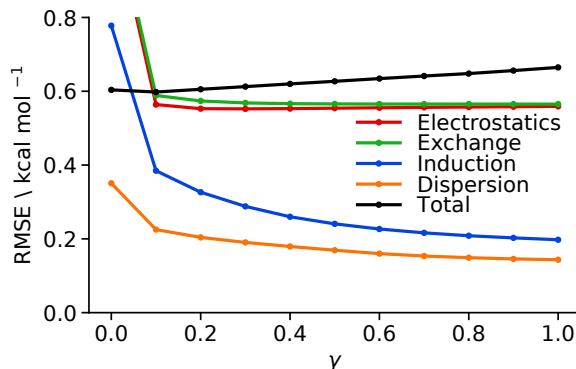


FIG. 3. Root mean squared errors in total and component energies (kcal mol^{-1}) as a function of γ , computed using our validation set. A value of $\gamma = 0$ corresponds to parameter fitting with respect to total energies only, and a value of $\gamma = 1.0$ corresponds to fitting purely to component energies.

t where $C = \{\text{electrostatics, exchange, induction, dispersion}\}$.⁵⁰ For a given value of γ , we minimize $\mathcal{L}$ using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm on the training set containing 7,884 dimers. We performed our parameter fitting in two stages. First, we fit the individual components to their corresponding SAPT component energy, constraining only that all parameters be nonnegative. We then refit all parameters simultaneously according to equation 22, using individually-optimized parameters as input. This two-step procedure can be seen as a refinement of the individually-optimized parameters, the extent of which being controlled by γ . To guide our choice of γ in determining a final parameter set, we use 5,000 dimers from the first collection of our custom-built database not involved in the fitting as a validation set. In Figure 3, we show the RMSE of the component and total energies of the validation set as function of γ . Aside from $\gamma = 0$, increasing values of γ tends to slightly decrease the quality of the total energy, while improving the component energies. Interestingly, the accuracy of each component seems to converge relatively quickly with γ , implying that choosing $\gamma = 1$ does not necessarily lead to a noticeable decrease in the quality of component energies, particularly for electrostatics and exchange-repulsion. To balance accuracy in component and total energies, we choose $\gamma = 0.4$ for determining our final parameters in the remainder of this work. Lastly, we show a summary of the computation of an interaction energy using CLIFF in Figure 4, highlighting how both the atomic parameters and the global parameters are used consistently in each component.

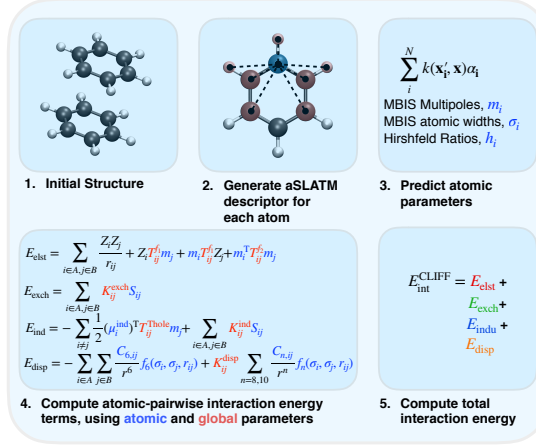


FIG. 4. Summary of interaction energy computation in CLIFF. Starting with a dimer (1), we generate all aSLATM descriptors for all atoms in each molecule (2). Using these descriptors, we then predict atomic MBIS multipoles, MBIS widths, and Hirshfeld ratios using KRR models (3). With these atomic properties, and our set of fitted global parameters, we compute each interaction energy term (4). Finally, we can compute the total interaction energy by summing the four components (5), analogously to SAPT.

III. RESULTS

A. Accuracy of Machine-Learned Atomic Parameters

We summarize mean absolute errors (MAE) of the atomic charges, dipoles, and quadrupoles computed from CLIFF with respect to those derived directly from MBIS densities in Table II. Correlation plots of the multipoles showing their magnitudes can be found in Figure S1 in the SI. The models from IPML report MAEs of $0.01 e$, $0.01 e\text{\AA}$, and $0.02 e\text{\AA}^2$ for charges, dipoles, and quadrupoles averaged across C, N, O, and H elements. Performing analogous averaging using all of our elements yields MAEs of $0.013 e$, $0.013 e\text{\AA}$, and $0.027 e\text{\AA}^2$. Our comparable accuracy with IPML is perhaps not surprising, as the only difference between our approaches is in our training data, implying that both models are near saturation. The larger errors seen for the multipoles of S are primarily a result of larger magnitudes of multipole moments, though it may be possible that these multipoles improve in accuracy with more training data. Additionally, larger errors in S, C, and N multipoles with respect to other atoms can also be attributed to the more diverse bonding environments those elements can have. Due to the high cost of KRR, we would have to use an alternative ML technique to make use of more data. Encouragingly, our errors for the halogens are

TABLE II. Mean absolute errors of predicted multipoles for each available element in CLIFF. Correlation plots comparing reference multipoles with ML multipoles can be found in the SI.

	C	N	O	H	S	F	Cl	Br
$q(e)$	0.019	0.015	0.009	0.005	0.024	0.009	0.012	0.011
$\mu(e\text{\AA})$	0.017	0.016	0.006	0.002	0.031	0.006	0.013	0.013
$Q(e\text{\AA}^2)$	0.019	0.024	0.009	0.003	0.131	0.005	0.012	0.012

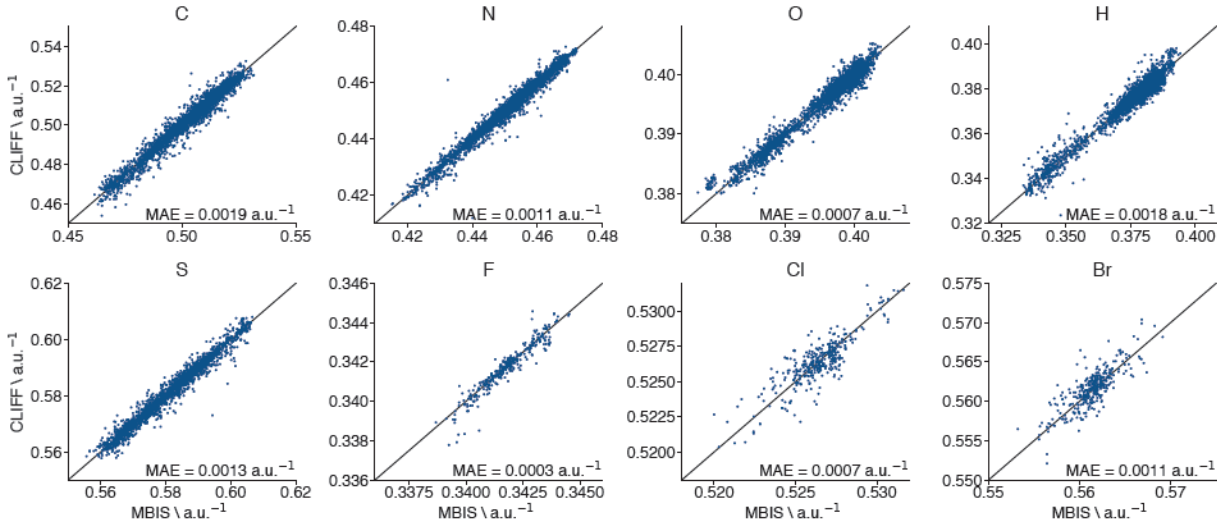


FIG. 5. Correlation of reference valence atomic widths computed using MBIS and ML widths computed in CLIFF.

very low, despite the relatively large magnitudes seen in the dipoles and quadrupoles.

Figures 5 and 6 show correlations between ML-predicted and reference atomic widths and Hirshfeld ratios, respectively, for all elements under study. Encouragingly, we see that for all elements, our ML models are able to accurately reproduce values computed from the *ab initio* densities. For the atomic widths, we see most of the errors to be around 0.001 a.u.^{-1} , with the largest being 0.0019 a.u.^{-1} for the carbon atom. Our averaged MAE of 0.0013 a.u.^{-1} , computed by averaging over all element types, represents a significant improvement over the similarly averaged MAE of 0.004 a.u.^{-1} from IPML. Such an improvement is likely a result of our use of the more sophisticated aSLATM descriptors, which are better suited to capture diverse bonding environments in comparison to Coulomb matrices. For the Hirshfeld ratios, our models produce a seemingly negligible improvement over those in IPML, with averaged MAEs of 0.0059 and

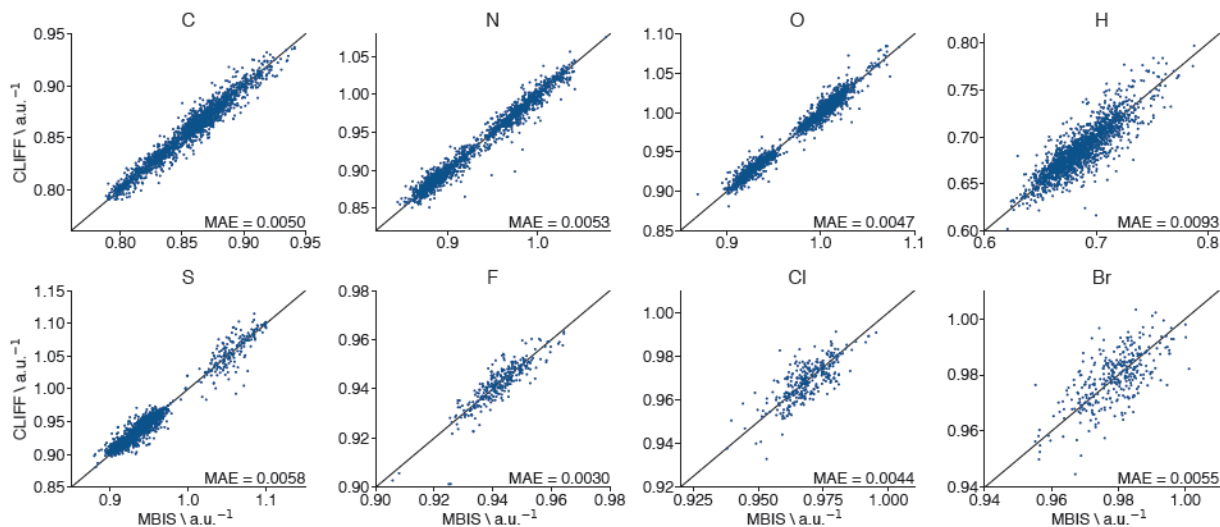


FIG. 6. Correlation of reference Hirshfeld ratios and machine-learned Hirshfeld ratios computed in CLIFF.

0.006, respectively. Importantly, our value does include S, Cl, F, and Br. Across all elements and all atomic properties, the high accuracy of our models can be attributed to a combination of better descriptors and larger training sets.

In general, the broad range of values for most element types supports our use of ML models for computing atom-environment dependent quantities rather than fitted parameters. In particular, the Hirshfeld ratios of S and O can be qualitatively different, with values being either above or below one, depending on the local environment. Conversely, the atomic widths and Hirshfeld ratios for the halogens all span a very narrow range of values. Perhaps unsurprisingly, this narrow window is a result of halogens having fairly uniform bonding tendencies resulting in similarly uniform atomic densities. Therefore, for these atoms and properties, one could be reasonably justified in using global parameters rather than ML models.

Finally, we note that KRR prediction of atomic properties in general is required for any energy computation. In a molecular dynamics simulation, for example, these atomic properties would be recomputed on-the-fly as the geometry of the molecule changes. Coupling atomic parameterization with computation of energy terms allows for potentially much more accurate simulations compared to using fixed parameters, assuming that the speed of parameterization is acceptably fast. Our choice of KRR results in predictions of atomic parameters that are far too slow to be used in routine dynamics simulations, though it does achieve improvement over quantum chemically derived parameters, particularly for large (>100 atom) systems. Fortunately, development

of much more efficient neural network-based atomic models could enable on-the-fly prediction of these atomic properties.

B. Small-Molecule Dimer Databases

To first understand the accuracy and transferability of CLIFF, we test against a variety of publicly-available, small-molecule dimer databases. Specifically, we use the S66x8,^{92–94} SSI,⁹⁵ X40x8,^{96,97} and NBCext10^{97–99} databases as they cover diverse interaction types including hydrogen bonding, halogen bonding, and interactions with one or more system often over radial potential energy surface scans. For X40x8, we removed dimers containing HCl, HBr, HF or iodine as we did not fit parameters for iodine or for those specific cases for hydrogen. Also, we retained only dimers that had inter-monomer separations between a factor of 0.90 and 2.0 of equilibrium, analogous to S66x8. Note that none of the dimers in these tests were used in fitting the global interaction parameters. We summarize MAEs, root-mean-squared errors (RMSEs), and maximum errors (MAX) for component and total interaction energies for all databases under study in Table III. All statistics are computed with respect to SAPT2+(3) MP2/aug-cc-pVTZ reference data. Across all databases, we achieve MAEs of ≤ 0.7 kcal mol⁻¹ or better for total and component interaction energies, with the exception of the exchange-repulsion energies of NBC10. CLIFF does particularly well for SSI and NBC10 total interaction energies, with errors of 0.28 and 0.49 kcal mol⁻¹, respectively, which we find to be particularly encouraging since these databases include diverse interaction types found in biological applications, including hydrogen bonding and a variety of π - π and π -H interactions. Lastly, the 0.614 kcal mol⁻¹ MAE in X40x8 also suggests that CLIFF does not perform significantly worse for computations on halogen-containing dimers.

We also plot the correlation of the total interaction energies predicted from CLIFF with the SAPT reference energies in Figure 7. For all databases, particularly those with radial scans, our errors grow significantly for dimers with positive (repulsive) interaction energies. Even for the SSI database, large errors are found for the few highly repulsive configurations shown at ~ 10 kcal mol⁻¹. Aside from the highly repulsive configurations, we see that CLIFF can reliably predict total interaction energies for a diverse set of interactions with quantitative accuracy. Additionally, we emphasize that all predictions are done using a single parameter set, and higher accuracy can be achieved in principle by using these test dimers in the fitting process. Such a procedure may be warranted for future applications of CLIFF on condensed systems with one or few different

TABLE III. MAE, RMSE, and maximum errors for component and total energies of various datasets computed with CLIFF. All statistics are computed with respect to SAPT2+(3) MP2/aug-cc-pVTZ reference values. All values are in kcal mol⁻¹.

		S66x8	SSI	X40x8	NBC10
electrostatics	MAE	0.507	0.328	0.467	0.650
	RMSE	0.822	0.481	0.905	1.176
	MAX	4.739	8.883	5.642	4.798
exchange	MAE	0.695	0.507	0.374	0.854
	RMSE	1.149	0.770	0.627	1.480
	MAX	5.844	13.954	2.725	6.124
induction	MAE	0.264	0.091	0.144	0.138
	RMSE	0.574	0.246	0.302	0.243
	MAX	4.734	3.928	1.633	1.171
dispersion	MAE	0.206	0.109	0.139	0.190
	RMSE	0.303	0.158	0.273	0.297
	MAX	1.381	1.664	1.518	0.924
total	MAE	0.703	0.281	0.614	0.492
	RMSE	1.115	0.547	1.176	0.939
	MAX	4.366	10.031	7.431	6.408

molecules. In general, we prefer optimization of a parameter set that is capable of treating many different interaction types reasonably well over one that is narrower in scope but with higher accuracy within that scope.

To understand some sources of error in our predictions of the total interaction energy for these databases, we first look at errors in our component energies. Both the induction and the dispersion components show very low errors consistent across all databases. In particular, our largest MAE for dispersion is only 0.206 kcal mol⁻¹ for the S66x8 database, followed by a 0.19 kcal mol⁻¹ for NBC10. Such small errors in dispersion for NBC10 is especially encouraging since this database

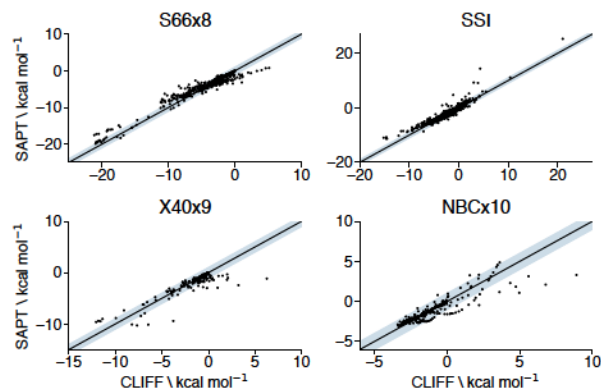


FIG. 7. Correlation between total interaction energies computed with CLIFF and with SAPT2+(3) δ MP2/aug-cc-pVTZ for a variety of test sets. The shaded blue area represents 1 kcal mol⁻¹ error.

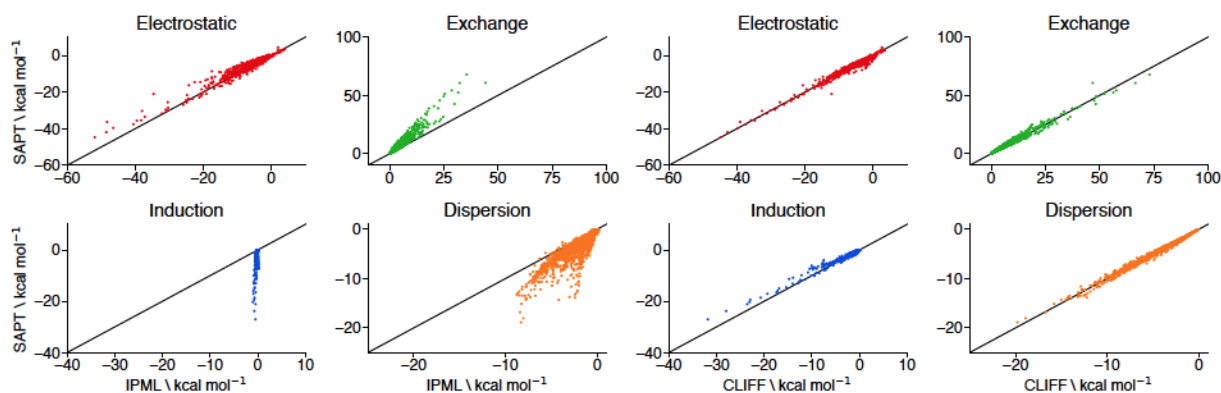


FIG. 8. Comparison of IPML and CLIFF components against SAPT2+(3) δ MP2/aug-cc-pVTZ reference values. Data obtained from a combination of S66x8, SSI, and NBC10 and databases where all dimers containing S are removed to enable direct comparison with IPML.

is composed mostly of challenging dispersion-bound complexes. In general, we see the exchange-repulsion energy usually serves as the error-limiting component, with electrostatics typically producing just slightly better performance. Not only are the MAEs for the exchange energies typically the largest, but they also can exhibit large outliers, with possible errors over 10 kcal mol⁻¹. These errors tend to be largest for short-range interactions that occur in repulsive regions of potential energy curves. At these ranges, the magnitude of exchange can easily surpass 50-60 kcal mol⁻¹, so that even errors of 10 kcal mol⁻¹ remain moderate in a relative sense.

In Figure 8, we plot correlations for component energies computed with IPML and CLIFF

TABLE IV. Total interaction energy MAEs and RMSEs computed using a variety of force fields, semi-empirical methods, and *s*SAPT0/jun-cc-pVDZ on the neutral subset of SSI with (All) and without (CNOH) sulfur-containing residues. All statistics are computed with respect to DW-CCSD(T**)-F12/aug-cc-pV(d+D)Z. All data aside from CLIFF and IPML values were taken from the BioFragment Database (BFDdb).⁹⁵ All values are in kcal mol⁻¹.

	CNOH		All	
	MAE	RMSE	MAE	RMSE
<i>s</i> SAPT0/jDZ	0.460	0.514	0.473	0.566
AM1	1.548	2.152	1.547	2.149
PM6-DH2	0.266	0.533	0.278	0.572
GAFF	0.620	1.627	0.620	1.598
CGenFF	0.308	1.060	0.319	1.066
IPML	0.387	0.680		
CLIFF	0.266	0.489	0.272	0.516

with respect to reference SAPT2+(3) MP2/aug-cc-pVTZ data. In the left half of the figure, we show the component predictions of IPML, and in the right half we show the same predictions from CLIFF, all on S66x8, NBC10, and SSI dimers. We leave out all S-containing molecules in this figure to enable comparisons with IPML, though those dimers are included in all other tables and figures. For these databases, we see that CLIFF shows very good correlation for all components, although the magnitudes of induction and exchange-repulsion are over- and under-predicted at large magnitudes, respectively. Despite this small degree of systematic error, in general the component energies are sufficiently accurate as to be useful for interpreting and classifying intermolecular interactions and remain minimally reliant on error cancellations for computing total interaction energies.

Parameters in IPML are fit to CCSD(T)/CBS total interaction energies, so its components are not guaranteed to match any SAPT data. Nonetheless, the electrostatic component in IPML contains no global parameters and does reasonably well, though tends to overpredict the magnitude of interactions, especially at shorter range. Interestingly, the induction fits nearly to zero for all dimers, and systematic deviations in exchange and dispersion are also present and rather large. The components in IPML cannot be used for qualitative analysis, but the total energies they pre-

dict are remarkably accurate for S66x8 and SSI, where they produce MAEs with respect to DW-CCSD(T**)-F12/aug-cc-pVTZ¹⁰⁰ of 0.4 kcal mol⁻¹ and 0.387 kcal mol⁻¹, respectively, using the S22 dataset for parameter fitting.⁵⁵ While very good, this result is in some ways surprising given that IPML predicts no significant induction, which is very important for hydrogen-bonded systems.

Both CLIFF and IPML perform favorably for total interaction energies of SSI in comparison to DW-CCSD(T**)-F12/aug-cc-pVTZ benchmark data. In Table IV, we compare total interaction energies of the neutral subset of SSI computed with CLIFF, IPML, and a variety of other force fields, semi-empirical methods, and SAPT0/jun-cc-pVDZ. The semi-empirical methods chosen are the common Austin Model 1 (AM1)¹⁰¹ and the Parameterized Model 6 with dispersion and hydrogen-bonding corrections (PM6-DH2).¹⁰² We also compare to the Generalized Amber Force Field (GAFF)¹⁰³ and the CHARMM General Force Field (CGenFF).^{104,105} Aside from IPML and CLIFF, all data were taken from the publicly-available BioFragment Database (BFDdb).⁹⁵ To enable comparisons with IPML, we present MAEs and RMSEs of the total energies of the subset of SSI containing only C, N, O, and H. For the remaining methods, we also compare the full neutral subset of SSI inclusive of sulfur-containing monomers. CLIFF performs remarkably well in comparison to all other methods, showing the lowest MAEs and RMSEs for both SSI sets. Such good agreement with the coupled-cluster reference suggests that our high-level SAPT reference data used to obtain global parameters is well-chosen for these neutral dimers. The performance of both CGenFF and GAFF in computing these gas phase interaction energies is also impressive, considering that both methods are fit to condensed phase properties rather than to interaction energies as is done with IPML and CLIFF. CGenFF and IPML both show low MAEs for these systems, but interestingly have fairly high RMSEs indicative of an increased presence of large errors in some dimers. Large errors typically occur when error cancellation is poor, and the low RMSE seen for CLIFF suggests that its accuracy is less dependent on cancellation of errors in comparison to the other methods shown. Interestingly, CLIFF, IPML, CGenFF, and PM6-DH2 all show somewhat lower MAEs compared to SAPT0/jDZ, although CLIFF is the only of these methods with a comparatively lower RMSE. Not only does CLIFF outperform all methods in this comparison, but, as previously shown, it does so with accurate component energies and without any required reparameterization and a modest number of distinct atom types.

TABLE V. MAEs and RMSEs for component and total energies of a combined S66x8, SSI, NBC10 and X40x8 dataset computed using atomic parameters from ML and AIM methods. All statistics are computed with respect to SAPT2+(3) MP2/aug-cc-pVTZ reference values. All values are in kcal mol⁻¹.

		ML	MBIS
electrostatics	MAE	0.389	0.416
	RMSE	0.653	0.759
exchange	MAE	0.555	0.504
	RMSE	0.904	0.859
induction	MAE	0.125	0.147
	RMSE	0.324	0.432
dispersion	MAE	0.133	0.128
	RMSE	0.209	0.203
total	MAE	0.384	0.444
	RMSE	0.748	0.940

C. Sensitivity of Functional Forms to Atomic Parameters

While prediction of component energies appears to be accurate, our use of ML models to generate atomic parameters may be introducing small errors. Errors in component energies come from three sources: inaccurate atomic parameters, overly-approximate functional forms, and poorly-obtained global parameters. While the accuracy of the reference AIM quantities we use to train our models remains somewhat of an open question, we can at least test the degree to which small errors introduced by the ML models affect the component energies. We use the term “AIM” here to refer to both MBIS and Hirshfeld routines used to generate reference atomic properties for CLIFF.

To empirically test the propagation of errors of our machine-learned atomic parameters, we compare errors in component and total energies for S66x8, SSI, NBC10, and X40x8 obtained from CLIFF using atomic parameters computed either with ML or directly from *ab initio* densities. In

Table V, we show MAEs, RMSEs, and MAX errors with respect to SAPT2+(3) MP2/aug-cc-pVTZ averaged over the four datasets, which were chosen so that all elements (C, N, O, H, S, F, Cl, Br) in CLIFF are represented. We see that both sources of atomic parameters give effectively identical statistics, with an absolute difference in total energy MAE of only 0.040 kcal mol⁻¹. Differences in component statistics are similarly small, with the largest difference being only 0.051 kcal mol⁻¹, seen in exchange-repulsion. Based on this data, we can effectively conclude that the reported errors in our ML models for atomic properties do not appear in a significant way in either our total or component interaction energies.

An important aspect of this experiment is that the standard CLIFF global parameters are used, and they are all fit using the ML atomic parameters. Therefore, it is reasonable to assume that there could be some bias in CLIFF that would favor using ML atomic parameters over the ones derived using AIM procedures directly from *ab initio* densities. Indeed, we do see that the MAEs and RMSEs are in general lower when using ML compared to AIM. This difference is very small, however, so we can confidently conclude that our ML atomic parameters are sufficiently accurate. Furthermore, the seeming interchangeability between ML and AIM parameters in CLIFF implies that our global parameters are not masking significant deficiencies in our ML parameters.

D. π - π and CH- π potential energy curves

Prediction of interaction energies for non-covalently bound dimers involving aromatic systems is a notoriously difficult problem for force fields.^{81,82,90,106,107} Not only do these types of dimers exhibit a typically strong attractive force from dispersion interactions, but charge penetration can also introduce a significant stabilizing effect at short range.⁷ Force fields that do not explicitly treat charge penetration thus rely on fortuitous error cancellations and overestimations of dispersion to obtain reasonable descriptions of π - π and CH- π interactions.^{107,108} CLIFF explicitly treats charge penetration using the damped multipole model introduced by Rackers et al.,⁶⁹ which has been shown to reproduce SAPT2+ electrostatic energies for benzene dimers and stacked DNA base pairs.⁶⁹ To further test the accuracy of CLIFF, and in particular the performance of our parameterization of the electrostatic functional form, we apply CLIFF to the sandwich, T-shaped, and parallel-displaced benzene dimer configurations in addition to the methane-benzene dimer. For each dimer, we compute potential energy curves using rigid monomers, and the parallel-displaced configuration has a constant vertical inter-monomer separation of 3.4 Å. All molecular

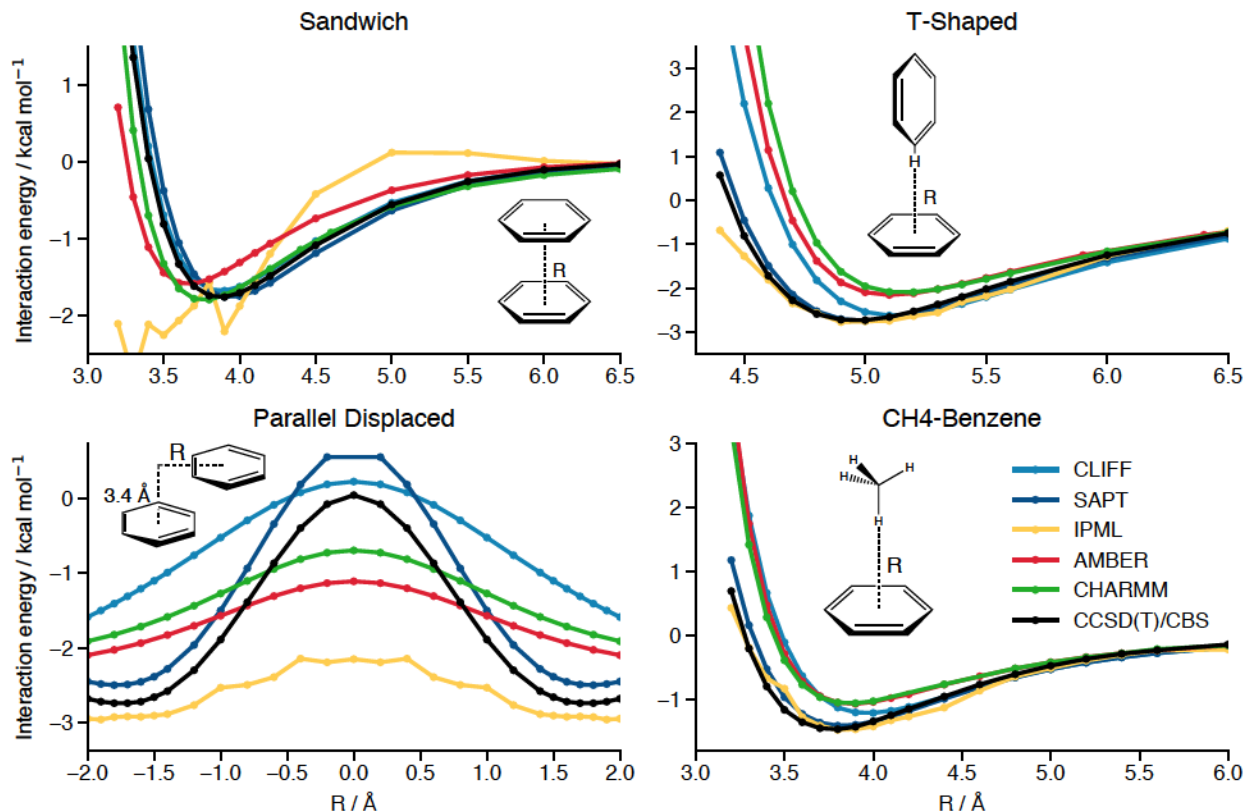


FIG. 9. Total interaction energies (kcal mol^{-1}) of the sandwich, T-shaped, and parallel displaced benzene dimer configurations and a methane-benzene complex all along a radial coordinate. We use a vertical displacement of $3.4 \text{ \AA}$ for the parallel displaced benzene dimer. In the legend, SAPT refers to $\text{SAPT2+(3)}\delta\text{MP2/aug-cc-pVTZ}$.

coordinates and reference CCSD(T)/CBS , AMBER, and CHARMM values are taken from Ref 107.

In Figure 9, we show the potential energy curves for the four dimers computed using CLIFF, $\text{SAPT2+(3)}\delta\text{MP2/aug-cc-pVTZ}$, IPML, AMBER, CHARMM, and CCSD(T)/CBS . Somewhat disappointingly, none of the empirical models, including CLIFF, treat these four systems with equal accuracy, despite explicit treatments of charge penetration in both CLIFF and IPML. In general, IPML shows noncontinuous potential energy curves that are qualitatively wrong for the sandwich and parallel-displaced configurations, though quite accurate for the T-shaped benzene dimer and the methane-benzene dimer. This non-continuity is a result of shifting local reference frames, and can in principle be cured with a different definition of local axes. AMBER and CHARMM perform reasonably well for the sandwich dimer, but they significantly overbind the

T-shaped and methane-benzene dimers. These non-polarizable force-fields also converge to the correct asymptotic limit rather slowly, where close agreement with CCSD(T)/CBS is seen only at very large intermolecular separations.

CLIFF shows good agreement with CCSD(T)/CBS and SAPT data near equilibrium intermolecular distances all the way through dissociation, though it underbinds the T-shaped and methane-benzene dimers. Since CLIFF is parameterized with respect to SAPT, agreement between CLIFF and CCSD(T)/CBS is only really possible when our chosen SAPT method itself agrees well with CCSD(T)/CBS. Based on previous analyses,⁹¹ this agreement is expected to be good for the organic, neutral molecules in this study. For the T-shaped and methane-benzene dimers in particular, CLIFF becomes far too repulsive at short range, similar to AMBER and CHARMM. This over-repulsion at short range predicted with CLIFF is consistent with the increased errors for repulsive configurations shown previously. While CLIFF performs very well even at short range for the sandwich configuration, it fails for the parallel-displaced configuration at a short interplanar distance of 3.4 Å. Interestingly, CLIFF, AMBER, and CHARMM all show a similarly broad potential for the parallel-displaced configuration, though the barrier height is somewhat more accurately predicted in CLIFF. For all four dimers, CLIFF on average matches reference CCSD(T)/CBS reference data better than the other empirical models shown, though in some cases CLIFF only does marginally better than much simpler non-polarizable models.

To identify sources of error in these systems, we compare component energies computed using CLIFF and SAPT2+(3) MP2/aug-cc-pVTZ for the sandwich, T-shaped, and parallel-displaced dimer configurations in Figure 10. For the sandwich dimer, good agreement in total interaction energies is achieved at short range by a cancellation of errors between an underestimated exchange energy and an overestimated (in an absolute sense) electrostatic energy. The non-systematic nature of the errors in exchange combined with the incorrect shape of the exchange component curve suggest that some errors related to our neglect of anisotropy may be present. Interestingly, the electrostatics in the T-shaped configuration agree quite well at all displacements, but the exchange energy is now overestimated, causing the observed overbinding, though this overestimation is relatively modest.

For the parallel-displaced configuration, electrostatics are predicted to be near-zero, suggesting an incomplete treatment of charge penetration effects. Additionally, the exchange repulsion is too small and too broad, causing a similarly broad total energy curve. Encouragingly, our induction and dispersion components agree very well with the SAPT reference for all configurations studied.

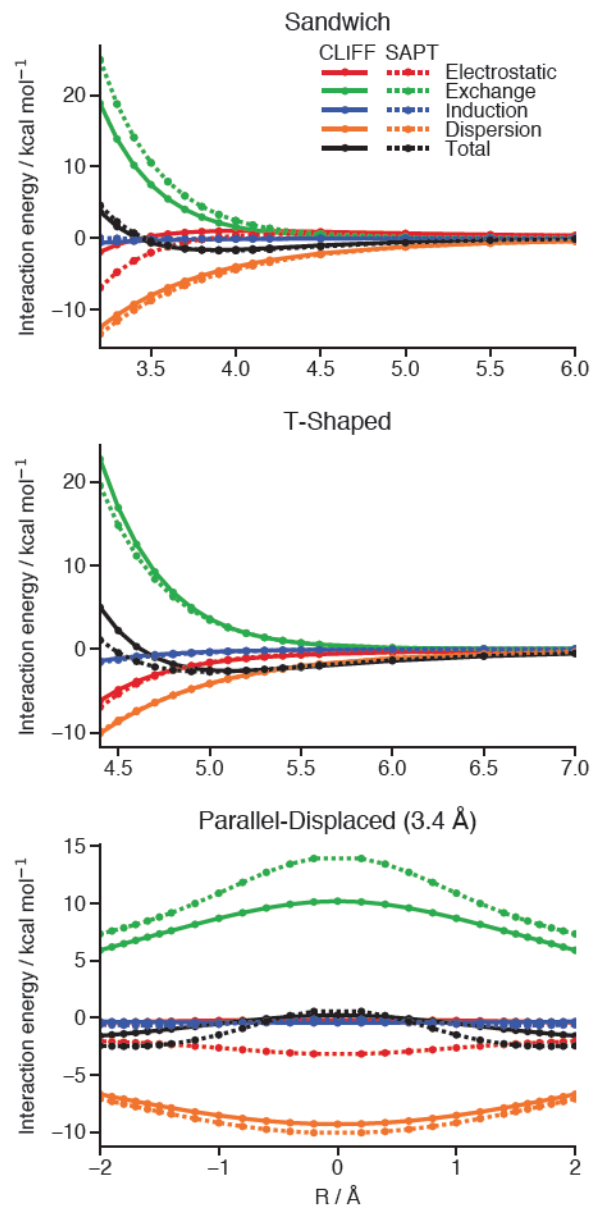


FIG. 10. Interaction energy components of sandwich, T-shaped, and parallel displaced benzene dimers computed with CLIFF (solid lines) and SAPT2+(3) δ MP2/aug-cc-pVTZ (dashed lines).

In previous work, the failure of AMBER and CHARMM to reproduce the parallel-displaced benzene dimer potentials was attributed to the lack of charge penetration in those models.¹⁰⁷ Specifically for sandwich and parallel-displaced benzene dimers, a pure multipole model will largely produce positive values for electrostatic interactions since the dominant energetic contributions come from interactions of like charges on carbon atoms and like charges on hydrogen atoms. CLIFF, alternatively, does explicitly treat charge penetration, and does so with high accuracy when the

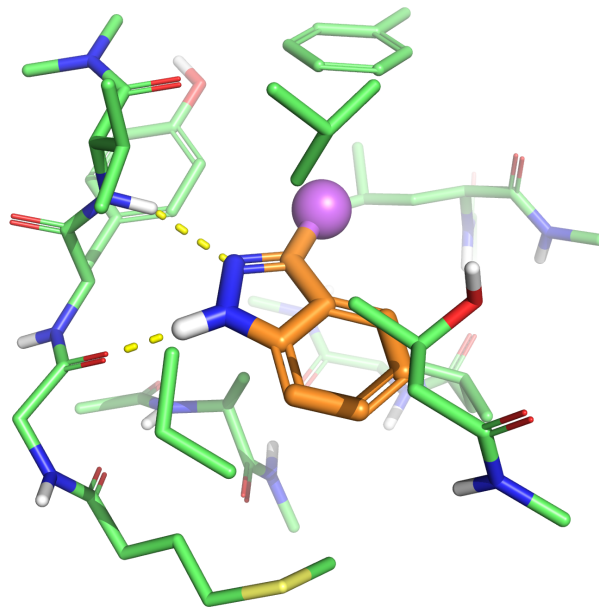


FIG. 11. Interaction between a chlorinated ligand fragment and a model binding pocket from the crystallographic structure 2UZU.

charge overlap is modest as seen with the T-shaped conformation. In cases of significant charge overlap, such as the sandwich conformation at very short range, our charge penetration model seems to more significantly underestimate stabilization effects. This underestimation points to some combination of deficiencies in the damping functions employed and the parameterizations used therein.

E. Protein-Ligand Interactions

As a final test, we use CLIFF to compute the interaction energies of a variety of protein-ligand complexes for which computation of SAPT0/jun-cc-pVDZ data is possible. To enable comparisons with SAPT0/jun-cc-pVDZ, we have generated a set of global parameters fit with respect to SAPT0/jun-cc-pVDZ computations on the same set of dimers used in higher-level CLIFF fits. Not only do these comparisons allow us to evaluate the performance of CLIFF on larger molecular systems, but it also lets us indirectly compare differences in low-level vs high-level SAPT when applied to large proteins. We will refer to these two parameterizations as CLIFF0 and CLIFF, where parameterizations come from SAPT0/jun-cc-pVDZ and SAPT2+(3) MP2/aug-cc-pVTZ, respectively.

TABLE VI. E computed for a set of five kinase, four fXa, and one LFA-1 binding pockets using SAPT0/jun-cc-pVDZ, CLIFF0, and CLIFF. We define E as the change in interaction energy when replacing a chloro- group with a methyl group on the ligand. We identify each protein pocket using the PDB label. All units are in kcal mol⁻¹.

System	SAPT0	CLIFF0	CLIFF
4YHT	1.89	2.01	2.27
2JDS	0.63	1.37	0.59
2OJF	1.36	2.09	1.17
2UZU	0.62	1.53	0.74
4YFF	0.21	0.46	0.07
2PR3	0.10	0.20	1.70
2W26	0.72	0.50	1.19
3ENS	0.61	1.32	0.10
2CJI	0.38	0.02	1.43
2O7N	2.79	3.65	2.91

Our protein-ligand test set is designed to test the ability of CLIFF and CLIFF0 to match reference SAPT0 data in predicting relative binding strengths between two similar ligands in the same binding region. This test set is made up of a variety of binding pockets from five kinases, four factor Xa (fXa) substructures, and one from a LFA-1 derived hydrophobic pocket. To each of the kinases and fXa substructures, we associate two ligands that differ only by substitution of a chlorine with a methyl group. For the LFA-1 system, the ligand substitution is between bromine and a cyano group. The protein-ligand systems contain between 144 and 228 atoms, and we show the 2UZU kinase binding pocket with chlorinated ligand as a representative example in Figure 11, which is used in this set. We then define the change in interaction energy, E as the difference in gas phase interaction energy between the methylated ligand and the chlorinated ligand for the kinase and fXa systems. For LFA-1, we define E as the difference in interaction energy between the cyanated ligand and the brominated ligand.

We report the E for each protein subsystem in Table VI, computed using SAPT0/jun-cc-pVDZ, CLIFF0, and CLIFF, and we label each system using the PDB ID from which each subsystem was derived. In all but one case, CLIFF0 is able to produce the same ligand ordering as

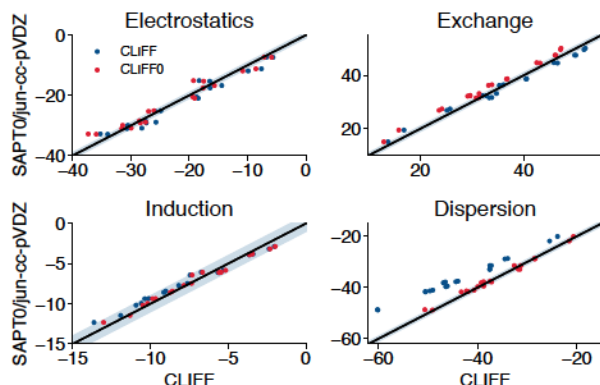


FIG. 12. Correlation between interaction energy components (kcal mol^{-1}) computed with SAPT0/jun-cc-pVDZ, CLIFF0, and CLIFF on the set of 20 protein-ligand complexes. A 1 kcal mol^{-1} error window is highlighted in blue.

SAPT0. Furthermore, CLIFF0 even shows some quantitative agreement with SAPT0, with half of the systems giving less than $0.3 \text{ kcal mol}^{-1}$ absolute error with respect to SAPT0, compared to a $0.6 \text{ kcal mol}^{-1}$ mean absolute error. The 3ENS fXa system shows the largest error of nearly 2 kcal mol^{-1} , and also results in CLIFF0 incorrectly predicting the preferred ligand. In general, however, the CLIFF0 results appear to be quite reliable in reproducing conclusions made using SAPT0 data. Unlike CLIFF0, the CLIFF results are not constructed to match those of SAPT0, and in certain cases significant disagreements arise. In 4YFF and 2W26, for example, CLIFF predicts the chlorinated ligands to have a more favorable interaction compared to the methylated ligands. For the latter case in particular, CLIFF predicts a $1.19 \text{ kcal mol}^{-1}$ stabilization upon chlorination, compared to the $0.72 \text{ kcal mol}^{-1}$ destabilization predicted by SAPT0. The observed reliability of CLIFF0 with respect to SAPT0 does imply that CLIFF, being fit with respect to SAPT2+(3) δ MP2/aug-cc-pVTZ, may be producing a more accurate result than SAPT0.

To benchmark the performance of CLIFF0 and CLIFF on these systems, we plot the correlation of their component energies in Figure 12. For CLIFF0, we see that the errors with respect to SAPT0 in induction and dispersion are largely within a 1 kcal mol^{-1} error window, and that electrostatics and exchange show slightly larger errors but still good correlation. For electrostatics, exchange, and induction, the CLIFF component energies also seem to correlate fairly well with SAPT0, suggesting that SAPT0/jun-cc-pVDZ and SAPT2+(3) δ MP2/aug-cc-pVTZ produce similar values for these components. For dispersion, however, the CLIFF predictions are systemat-

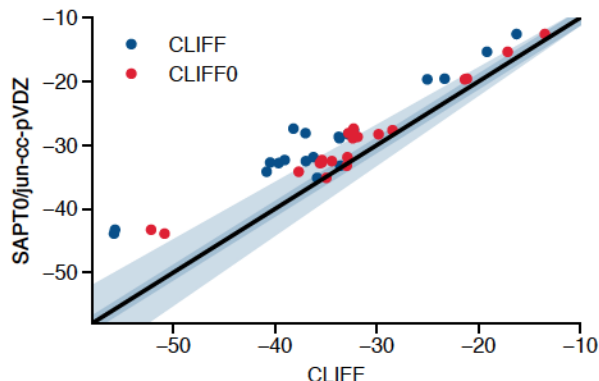


FIG. 13. Correlation between total interaction energies (kcal mol^{-1}) computed with SAPT0/jun-cc-pVDZ, CLIFF0, and CLIFF on the set of 20 protein-ligand complexes. A 1 kcal mol^{-1} error window is highlighted in dark blue, and a 10% error window is highlighted in light blue.

ically larger in magnitude compared to both CLIFF0 and SAPT0, with the largest difference being over 10 kcal mol^{-1} .

CLIFF’s prediction of higher-magnitude dispersion compared to SAPT0 and CLIFF0 also results in total interaction energies being similarly larger in magnitude, as shown in Figure 13. Across a fairly broad energy range, CLIFF0 usually produces less than 10% error compared to SAPT0 interaction energies, and at times shows less than a 1 kcal mol^{-1} error. CLIFF predicts more stable protein-ligand interactions compared to SAPT0 and CLIFF0, consistent with the larger predicted contributions from dispersion interactions. Based on the demonstrated accuracy of CLIFF0 and CLIFF, it is likely that SAPT0 underbinds these complexes to some extent due to neglected dispersion interactions. Additionally, the underestimated dispersion in SAPT0 is not uniform across this test set, implying that it may not always be reliable in rank-ordering ligands.

IV. CONCLUSIONS

We have presented a new intermolecular force field designed to have transferable accuracy in computing both component and total interaction energies for diverse interaction types encountered in organic chemistry and drug discovery applications. Our model, CLIFF, achieves this transferrable accuracy through a combination of machine learning and advanced functional forms from symmetry adapted perturbation theory (SAPT) based force fields. The transferability in CLIFF is achieved from its adoption of functional forms that traditionally use quantities derived

from molecular densities to reduce the amount of parameter fitting required. Following IPML, the applicability of CLIFF is further increased by machine-learning these atomic parameters, so that computation of any new, potentially large molecule can be done nearly automatically. Lastly, good accuracy for many diverse large and small molecules is achieved by using advanced functional forms from SAPT-based force fields.^{26,28,69} Furthermore, we reliably obtain accurate component and total energies by fitting our modest number of global parameters to high-level SAPT data.

We tested CLIFF on diverse small-molecule databases, where we achieved MAEs of at worst 0.70 kcal mol⁻¹ for the radial scans of 66 van der Waals dimers in the S66x8 test set, and at best 0.28 kcal mol⁻¹ for the neutral-neutral interaction energies of the sidechain-sidechain interactions (SSI) test set, derived from close contacts in the protein databank (PDB). For short-range, repulsive configurations, CLIFF produced its largest errors, and tended to overbind molecules in this energetic region, at times substantially. From equilibrium intermolecular distances to dissociation, conversely, CLIFF matches reference high-level quantum mechanical SAPT interaction total and component energies with consistently low errors. In applications to ligands in protein pockets, where structures are usually desired in an optimized configuration, CLIFF is thus expected to perform favorably, as these systems require accurate computation of numerous long-range interactions in addition to a smaller number of dominant contacts. Our tests on a variety of 20 protein-ligand interactions support this notion, where CLIFF0 was able to reliably reproduce SAPT0 reference data. Moreover CLIFF’s parameterization to higher-level SAPT makes it a very useful tool in studying these systems, in that it provides potentially more accuracy than SAPT0. Lastly, CLIFF represents one of the only ways to routinely achieve the accuracy of *ab initio* density-based, SAPT-based force fields for systems containing hundreds of atoms without requiring massive *ab initio* computations intractable for many applications.

Despite these initial successes of CLIFF, several challenges persist that need to be addressed in future developments. Perhaps the most significant limitation is that CLIFF, like IPML, can only be applied to neutral dimers. Extension of CLIFF to charged systems would require nontrivial changes to the parameterization scheme, formulation of machine learning models, and potentially even the functional forms. While our reference SAPT data is reliable for the systems in this work, it is unclear if the same reliability of the reference data can be expected for charged dimers. Additionally, our application of CLIFF to dimers involving benzene in Section III.D highlighted a few deficiencies in the current model. While our model for electrostatics is shown to be accurate in the larger databases, we see it can underestimate the extent of charge penetration in certain

cases. From these same tests, we also saw that our model for exchange-repulsion can at times be inaccurate, likely due to our neglect of anisotropy in our overlap-based formulation.

CLIFF, like IPML, represents an intermediate approach between pure ML potentials and *ab initio* parameterized force fields. Instead of using pure ML approaches to compute energy components directly, we use the advanced physics-based functional forms typical of *ab initio* force fields. However, instead of labor-intensive parameterization, or parameterization based on fragment quantum mechanical computations, we obtain atomic parameters by ML models that need only monomer geometries. Due to the increasing accuracy of ML potentials,⁵¹ however, a future direction of force field development in CLIFF may be to include machine-learned interaction energy models for some components, particularly for cases where the existing functional forms are least effective. Alternatively, improvements to the physical functional forms would also significantly help the model, particularly corrections for anisotropy.²⁷ Aside from improvements to CLIFF that result in higher accuracy, we also see a number of practical directions for CLIFF development. Most notably are the replacement of the KRR models with much faster neural-network based models, and adaptation of such models to enable prediction on charged molecules. With the speed of neural-networks, existing intramolecular force fields could use CLIFF as an intermolecular potential in molecular dynamics simulations to compute properties beyond interaction energies.

SUPPLEMENTARY MATERIAL

The supplementary material includes additional structural analyses on both monomer and dimer databases. We also report performance of our ML models in predicting atomic multipoles, global parameters used for CLIFF0, and all interaction energies used in the energy scans of the benzene dimer configurations and the benzene-methane complex. Total and component interaction energies for all protein-ligand complexes are also provided from SAPT0/jun-cc-pVDZ, CLIFF, and CLIFF0 computations. We include geometries of all protein-ligand complexes, all dimers used in fitting global parameters, and all monomers used in ML model training. Finally, reference interaction energies for our dimer database and reference atomic properties for the monomer database are also provided.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from Bristol Myers Squibb. J.B.S. also thanks Dr. Mary Van Vleet for very helpful discussions and comments regarding the manuscript and functional forms used in CLIFF.

DATA AVAILABILITY

The data that supports the findings of this study are available within the article and its supplementary material. Our reference implementation of CLIFF is publicly available on Github at <https://github.com/jeffschriber/cliff>.

REFERENCES

- ¹C. D. Sherrill, T. Takatani, and E. G. Hohenstein, *J. Phys. Chem. A* **113**, 10146 (2009).
- ²S. Grimme, *WIREs Comput. Mol. Sci.* **1**, 211 (2011).
- ³E. G. Hohenstein and C. D. Sherrill, *WIREs Comput. Mol. Sci.* **2**, 304 (2012).
- ⁴K. Szalewicz, *WIREs Comput Mol Sci* **2**, 254 (2012).
- ⁵K. Patkowski, *WIREs Comput Mol Sci* **10**, 1442 (2020).
- ⁶E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, J. M. Turney, and H. F. Schaefer, *J. Chem. Phys.* **135**, 174107 (2011).
- ⁷C. D. Sherrill, *Acc. Chem. Res.* **46**, 1020 (2013).
- ⁸R. M. Parrish, D. F. Sitkoff, D. L. Cheney, and C. D. Sherrill, *Chemistry* **23**, 7887 (2017).
- ⁹A. J. Stone and A. J. Misquitta, *International Reviews in Physical Chemistry* **26**, 193 (2007).
- ¹⁰J. G. McDaniel and J. R. Schmidt, *J. Phys. Chem. A* **117**, 2053 (2013).
- ¹¹J. G. McDaniel and J. R. Schmidt, *Annu. Rev. Phys. Chem.* **67**, 467 (2016).
- ¹²M. M. Ghahremanpour, P. J. van Maaren, C. Caleman, G. R. Hutchison, and D. van der Spoel, *J. Chem. Theory Comput.* **14**, 5553 (2018).
- ¹³K. Yu, J. G. McDaniel, and J. R. Schmidt, *J. Phys. Chem. B* **115**, 10054 (2011).
- ¹⁴J. R. Schmidt, K. Yu, and J. G. McDaniel, *Acc. Chem. Res.* **48**, 548 (2015).
- ¹⁵J. G. McDaniel, C. Y. Son, and A. Yethiraj, *J. Phys. Chem. B* **122**, 4101 (2018).
- ¹⁶K.-j. Jeong, J. G. McDaniel, and A. Yethiraj, *J. Phys. Chem. B* **124**, 7475 (2020).
- ¹⁷A. J. Misquitta, A. J. Stone, and F. Fazeli, *J. Chem. Theory Comput.* **10**, 5405 (2014).

- ¹⁸S. Grimme, J. Chem. Theory Comput. **10**, 4497 (2014).
- ¹⁹S. Grimme, C. Bannwarth, E. Caldeweyher, J. Pisarek, and A. Hansen, J. Chem. Phys. **147**, 161708 (2017).
- ²⁰D. J. Cole, J. Z. Vilseck, J. Tirado-Rives, M. C. Payne, and W. L. Jorgensen, J. Chem. Theory Comput. **12**, 2312 (2016).
- ²¹A. J. Stone, Chem. Phys. Lett. **83**, 233 (1981).
- ²²A. J. Stone, J. Chem. Theory Comput. **1**, 1128 (2005).
- ²³N. Rai and J. I. Siepmann, J. Phys. Chem. B **117**, 273 (2012).
- ²⁴A. J. Misquitta and A. J. Stone, J. Chem. Theory Comput. **12**, 4184 (2016).
- ²⁵A. J. Misquitta and A. J. Stone, Theor. Chem. Acc. **137**, 1 (2018).
- ²⁶M. J. Van Vleet, A. J. Misquitta, A. J. Stone, and J. R. Schmidt, J. Chem. Theor. Comput. **12**, 3851 (2016).
- ²⁷M. J. Van Vleet, A. J. Misquitta, and J. R. Schmidt, J. Chem. Theor. Comput. **14**, 739 (2018).
- ²⁸S. Vandenbrande, M. Waroquier, V. V. Speybroeck, and T. Verstraelen, J. Chem. Theory Comput. **13**, 161 (2017).
- ²⁹P. Ren, C. Wu, and J. W. Ponder, J. Chem. Theory Comput. **7**, 3143 (2011).
- ³⁰Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, J. Chem. Theory Comput. **9**, 4046 (2013).
- ³¹H. Chu, X. Peng, Y. Li, Y. Zhang, and G. Li, Molecules **23**, 77 (2017).
- ³²C. Zhang, D. Bell, M. Harger, and P. Ren, J. Chem. Theory Comput. **13**, 666 (2017).
- ³³C. Liu, J.-P. Piquemal, and P. Ren, J. Chem. Theory Comput. **15**, 4122 (2019).
- ³⁴M. L. Laury, Z. Wang, A. S. Gordon, and J. W. Ponder, J. Comput. Aided Mol. Des. **32**, 1087 (2018).
- ³⁵P. V. G. M. Rathnayake, S. Bernardi, and A. Widmer-Cooper, J. Chem. Phys. **152**, 024117 (2020).
- ³⁶K. Yao, J. E. Herr, D. W. Toth, R. Mckintyre, and J. Parkhill, Chem. Sci. **9**, 2261 (2018).
- ³⁷P. Gkeka, G. Stoltz, A. Barati Farimani, Z. Belkacemi, M. Ceriotti, J. D. Chodera, A. R. Dinner, A. L. Ferguson, J.-B. Maillet, H. Minoux, C. Peter, F. Pietrucci, A. Silveira, A. Tkatchenko, Z. Trstanova, R. Wiewiora, and T. Lelièvre, J Chem. Theory Comput. (2020).
- ³⁸D. Koner, S. M. Salehi, P. Mondal, and M. Meuwly, J. Chem. Phys. **153**, 010901 (2020).
- ³⁹J. Pei, L. F. Song, and K. M. Merz Jr., J. Chem. Theory Comput. **16**, 5385 (2020).
- ⁴⁰J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).

- ⁴¹J. Behler, J. Chem. Phys. **134**, 074106 (2011).
- ⁴²J. S. Smith, O. Isayev, and A. E. Roitberg, Chem. Sci. **8**, 3192 (2017).
- ⁴³K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, Nat. Comm. **8**, 13890 (2017).
- ⁴⁴N. Lubbers, J. S. Smith, and K. Barros, J. Chem. Phys. **148**, 241715 (2018).
- ⁴⁵K. T. Schütt, H. E. Saucedo, P. J. Kindermans, A. Tkatchenko, and K. R. Müller, J. Chem. Phys. **148**, 241722 (2018).
- ⁴⁶J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev, and A. E. Roitberg, Nat. Comm. **10**, 2903 (2019).
- ⁴⁷O. T. Unke and M. Meuwly, J. Chem. Theory Comput. **15**, 3678 (2019).
- ⁴⁸A. S. Christensen, L. A. Bratholm, F. A. Faber, and O. A. von Lilienfeld, J. Chem. Phys. **152**, 044107 (2020).
- ⁴⁹C. Devereux, J. S. Smith, K. K. Davis, K. Barros, R. Zubatyuk, O. Isayev, and A. E. Roitberg, J. Chem. Theory Comput. **16**, 4192 (2020).
- ⁵⁰D. P. Metcalf, A. Koutsoukas, S. A. Spronk, B. L. Claus, D. A. Loughney, S. R. Johnson, D. L. Cheney, and C. D. Sherrill, J. Chem. Phys. **152**, 074103 (2020).
- ⁵¹Z. L. Glick, D. P. Metcalf, A. Koutsoukas, S. A. Spronk, D. L. Cheney, and C. D. Sherrill, J. Chem. Phys. **153**, 044112 (2020).
- ⁵²J. Wang, P. Cieplak, J. Li, Q. Cai, M.-J. Hsieh, R. Luo, and Y. Duan, J. Phys. Chem. B **116**, 7088 (2012).
- ⁵³Y. Li, H. Li, F. C. Pickard IV, B. Narayanan, F. G. Sen, M. K. Y. Chan, S. K. R. S. Sankaranarayanan, B. R. Brooks, and B. Roux, J. Chem. Theory Comput. **13**, 4492 (2017).
- ⁵⁴R. Galvelis, S. Doerr, J. M. Damas, M. J. Harvey, and G. De Fabritiis, J. Chem. Inf. Model **59**, 3485 (2019).
- ⁵⁵T. Bereau, R. A. DiStasio Jr, A. Tkatchenko, and O. A. von Lilienfeld, J. Chem. Phys. **148**, 241706 (2018).
- ⁵⁶R. F. W. Bader, Chem. Rev. **91**, 893 (1991).
- ⁵⁷E. Pastorczak and C. Corminboeuf, J. Chem. Phys. **146**, 120901 (2017).
- ⁵⁸T. Verstraelen, S. Vandenbrande, F. Heidar-Zadeh, L. Vanduyfhuys, V. Van Speybroeck, M. Waroquier, and P. W. Ayers, J. Chem. Theory Comput. **12**, 3894 (2016).
- ⁵⁹F. L. Hirshfeld, Theoret. Chim. Acta **44**, 129 (1977).
- ⁶⁰A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).

- ⁶¹M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, *Phys. Rev. Lett.* **108**, 058301 (2012).
- ⁶²B. Huang and O. A. von Lilienfeld, *J. Chem. Phys.* **145**, 161102 (2016).
- ⁶³B. Huang and O. A. von Lilienfeld, *Nat. Chem.* **12**, 945 (2020).
- ⁶⁴A. S. Christensen, A. Faber, B. Huang, L. A. Bratholm, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, *QML: A python toolkit for quantum machine learning* (2017).
- ⁶⁵D. Mendez, A. Gaulton, A. P. Bento, J. Chambers, M. De Veij, E. Félix, M. P. Magariños, J. F. Mosquera, P. Mutowo, M. Nowotka, M. Gordillo-Marañón, F. Hunter, L. Junco, G. Mugumbate, M. Rodriguez-Lopez, F. Atkinson, N. Bosc, C. J. Radoux, A. Segura-Cabrera, A. Hersey, and A. R. Leach, *Nucleic Acids Res.* **47**, D930 (2019).
- ⁶⁶D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz, M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O'Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, H. F. Schaefer, A. Y. Sokolov, K. Patkowski, A. E. DePrince, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, *J Chem Phys* **152**, 184108 (2020).
- ⁶⁷T. Verstraelen, P. Tecmer, F. Heidar-Zadeh, C. E. González-Espinoza, M. Chan, T. D. Kim, K. Boguslawski, S. Fias, S. Vandenbrande, D. Berrocal, and P. W. Ayers, *HORTON 2.1.1* (2017).
- ⁶⁸Stone, Anthony J, *The Theory of Intermolecular Forces* (Oxford University Press, 2013).
- ⁶⁹J. A. Rackers, Q. Wang, C. Liu, J.-P. Piquemal, P. Ren, and J. W. Ponder, *Phys. Chem. Chem. Phys.* **19**, 276 (2017).
- ⁷⁰Y. S. Kim, S. K. Kim, and W. D. Lee, *Chem. Phys. Lett.* **80**, 574 (1981).
- ⁷¹C. Nyeland and J. P. Toennies, *Chem. Phys. Lett.* **127**, 172 (1986).
- ⁷²G. Ihm, M. Cole, F. Toigo, and J. Klein, *Phys. Rev. A* **42**, 5244 (1990).
- ⁷³N. Gresh, G. A. Cisneros, T. A. Darden, and J.-P. Piquemal, *J. Chem. Theory Comput.* **3**, 1960 (2007).
- ⁷⁴M. S. Gordon, L. Slipchenko, H. Li, and J. H. Jensen (Elsevier, 2007) pp. 177 – 193.
- ⁷⁵M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, *Chem. Rev.* **112**, 632 (2012).
- ⁷⁶M. S. Gordon, Q. A. Smith, P. Xu, and L. V. Slipchenko, *Annu. Rev. Phys. Chem.* **64**, 553 (2013).
- ⁷⁷A. J. Misquitta, B. Jeziorski, and K. Szalewicz, *Phys. Rev. Lett.* **91**, 033201 (2003).

- ⁷⁸A. J. Misquitta, R. Podeszwa, B. Jeziorski, and K. Szalewicz, *J. Chem. Phys.* **123**, 214103 (2005).
- ⁷⁹S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
- ⁸⁰E. Caldeweyher, C. Bannwarth, and S. Grimme, *J. Chem. Phys.* **147**, 034112 (2017).
- ⁸¹G. Prampolini, P. R. Livotto, and I. Cacelli, *J. Chem. Theory Comput.* **11**, 5182 (2015).
- ⁸²K. Strutyński, J. A. N. F. Gomes, and M. Melle-Franco, *J. Phys. Chem. A* **118**, 9561 (2014).
- ⁸³Y. Liu and I. William A Goddard, *J. Phys. Chem. Lett.* **1**, 2550 (2010).
- ⁸⁴K. U. Lao and J. M. Herbert, *J. Phys. Chem. Lett.* **3**, 3241 (2012).
- ⁸⁵M. Kim, W. J. Kim, T. Gould, E. K. Lee, S. Lebègue, and H. Kim, *J. A. C. S.* **142**, 2346 (2020).
- ⁸⁶K. T. Tang and J. P. Toennies, *J. Chem. Phys.* **80**, 3726 (1998).
- ⁸⁷X. Chu and A. Dalgarno, *J. Chem. Phys.* **121**, 4083 (2004).
- ⁸⁸B. Thole, *Chem. Phys.* **59**, 341 (1981).
- ⁸⁹P. Y. Ren and J. W. Ponder, *J. Phys. Chem. B* **107**, 5933 (2003).
- ⁹⁰T. S. Totton, A. J. Misquitta, and M. Kraft, *J. Chem. Theory Comput.* **6**, 683 (2010).
- ⁹¹T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, and C. D. Sherrill, *J. Chem. Phys.* **140**, 094106 (2014).
- ⁹²J. Rezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **7**, 2427 (2011).
- ⁹³J. Rezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **7**, 3466 (2011).
- ⁹⁴J. Rezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **10**, 1359 (2014).
- ⁹⁵L. A. Burns, J. C. Faver, Z. Zheng, M. S. Marshall, D. G. A. Smith, K. Vanommeslaeghe, A. D. MacKerell Jr., K. M. Merz Jr., and C. D. Sherrill, *J. Chem. Phys.* **147**, 161727 (2017).
- ⁹⁶J. Rezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **8**, 4285 (2012).
- ⁹⁷D. G. A. Smith, L. A. Burns, K. Patkowski, and C. D. Sherrill, *J. Phys. Chem. Lett.* **7**, 2197 (2016).
- ⁹⁸L. A. Burns, A. Vázquez-Mayagoitia, B. G. Sumpter, and C. D. Sherrill, *J. Chem. Phys.* **134**, 084107 (2011).
- ⁹⁹M. S. Marshall, L. A. Burns, and C. D. Sherrill, *J. Chem. Phys.* **135**, 194102 (2011).
- ¹⁰⁰M. S. Marshall and C. D. Sherrill, *J. Chem. Theory Comput.* **7**, 3978 (2011).
- ¹⁰¹M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. A. C. S.* **107**, 3902 (2002).
- ¹⁰²M. Korth, M. Pitonák, J. Rezáč, and P. Hobza, *J. Chem. Theory Comput.* **6**, 344 (2010).
- ¹⁰³D. A. Case, T. A. Darden, T. E. C. III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W.

- Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. Kollman, University of California, San Francisco (2012).
- ¹⁰⁴K. Vanommeslaeghe, E. P. Raman, and A. D. MacKerell, *J Chem Inf Model* **52**, 3155 (2012).
- ¹⁰⁵B. R. Brooks, C. L. Brooks, A. D. MacKerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Ca isch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, *J. Comp. Chem.* **30**, 1545 (2009).
- ¹⁰⁶A. T. Macias and A. D. Mackerell, *J. Comp. Chem.* **26**, 1452 (2005).
- ¹⁰⁷C. D. Sherrill, B. G. Sumpter, M. O. Sinnokrot, M. S. Marshall, E. G. Hohenstein, R. C. Walker, and I. R. Gould, *J. Comp. Chem.* **30**, 2187 (2009).
- ¹⁰⁸T. M. Parker and C. D. Sherrill, *J. Chem. Theory Comput.* **11**, 4197 (2015).