Evaluating Force-Field London Dispersion Coefficients Using the Exchange-Hole Dipole Moment Model

Mohamad Mohebifar,[†] Erin R. Johnson,[‡] and Christopher N. Rowley[*,†]

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

[‡]*Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, Canada*

E-mail: crowley@mun.ca

Abstract

London dispersion interactions play an integral role in materials science and biophysics. Force fields for atomistic molecular simulations typically represent dispersion interactions by the 12-6 Lennard-Jones potential, using empirically-determined parameters. These parameters are generally underdetermined and there is no straightforward way to test if they are physically realistic. Alternatively, the exchange-hole dipole moment (XDM) model from density-functional theory predicts atomic and molecular London dispersion coefficients from first principles, providing an innovative strategy to validate the dispersion terms of molecular-mechanical force fields. In this work, the XDM model was used to obtain the London dispersion coefficients of 88 organic molecules relevant to biochemistry and pharmaceutical chemistry and the values compared with those derived from the Lennard-Jones parameters of the CGenFF, GAFF, OPLS, and Drude polarizable force fields. The molecular dispersion coefficients for the CGenFF, GAFF, and OPLS models are systematically higher than the XDM-calculated

values by a factor of roughly 1.5, likely due to neglect of higher-order dispersion terms and premature truncation of the dispersion-energy summation. The XDM dispersion coefficients span a large range for some molecular-mechanical atom types, suggesting an unrecognized source of error in force-field models, which assume that atoms of the same type have the same dispersion interactions. Agreement with the XDM dispersion coefficients is even poorer for the Drude polarizable force field. Popular water models were also examined and TIP3P was found to have dispersion coefficients similar to the experimental and XDM references, although other models employ anomalously high values. Finally, XDM-derived dispersion coefficients were used to parameterize molecular-mechanical force fields for five liquids – benzene, toluene, cyclohexane, n-pentane, and n-hexane – which resulted in improved accuracy in the computed enthalpies of vaporization despite only having to evaluate a much smaller section of the parameter space.

1 Introduction

London dispersion interactions[1] arise from instantaneous dipole moments in the electron distribution of separated atoms or molecules, creating a universally attractive force between them. Although these interactions are generally weak, they play an integral role in the structure and dynamics of condensed matter due to their ubiquitous nature. In the field of biophysics, London dispersion is a critical element in lipid structure, membrane permeation,[2–5] protein folding,[6] protein–ligand binding,[7,8] and nucleic acid structure.[9] In materials science, London dispersion contributes to lattice energies,[10,11] crystal packing,[12–14] and surface adsorption.[15,16]

The potential energy of the London dispersion interaction between atoms i and j can be expressed as a multipolar expansion,[17]

$$E_{disp}(r_{ij}) = -\sum_{n=6}^{\infty} \frac{C_{n,ij}}{r_{ij}^{n}}, \tag{1}$$

where r_{ij} is the inter-atomic distance and $C_{n,ij}$ are the coefficients of the London dispersion interaction between atoms i and j for the nth order term of the interaction. In practice, odd-powered terms are generally negligible and the series is truncated at $n = 10$, yielding

$$E_{disp}(r_{ij}) = -\frac{C_{6,ij}}{r_{ij}^6} - \frac{C_{8,ij}}{r_{ij}^8} - \frac{C_{10,ij}}{r_{ij}^{10}} \tag{2}$$

where $C_{6,ij}$, $C_{8,ij}$, and $C_{10,ij}$ are the dispersion coefficients. These coefficients depend on the identity of the pair of interacting atoms. Various methods have been developed to infer these coefficients from experiment or to calculate them from first principles.[17,18] Ab initio-derived dispersion coefficients have been calculated by Cole et al.[19] and Vandenbrande et al.[20] using the Tkatchenko–Scheffler method, where the dispersion coefficients of free atoms were calculated using TDDFT then scaled according to the relative free volume of the atom in the molecule.[19,20]

One straightforward and non-empirical method for calculating the dispersion coefficients of atoms in molecules is the exchange-hole dipole moment (XDM) model.[21,22] The dispersion coefficients can be defined in terms of expectation values of the square of the instantaneous dipole moments resulting from a reference electron and its exchange hole, $\langle d_X^2 \rangle$, and the atom-in-molecule polarizabilities, α.

$$C_{6,ij} = \alpha_i \alpha_j \frac{\langle d_X^2 \rangle_i \langle d_X^2 \rangle_j}{\alpha_j \langle d_X^2 \rangle_i + \alpha_i \langle d_X^2 \rangle_j} \tag{3}$$

Because the dispersion coefficients are calculated directly from properties of the electron density, they vary with the local chemical environment of each atom within a molecule or solid.[15,16,23] Higher-order C_8 and C_{10} dispersion coefficients can similarly be obtained in terms of higher-order exchange-hole moment integrals.[21,22]

The molecular C_6 coefficient can be expressed simply as a sum of the atomic dispersion

coefficients for all pairs of atoms within the molecule:

$$C_{6,mol} = \sum_{ij} C_{6,ij}. \tag{4}$$

The XDM method was found to provide molecular C_6 dispersion coefficients that are in good agreement with the experimental values.[24] Evaluation of the XDM dispersion coefficients and dispersion energy can be performed routinely from a ground-state density-functional theory (DFT) calculation using standard quantum-chemistry codes.[22,25,26] The calculation of atomic London dispersion coefficients using XDM could provide an innovative method to validate force-field dispersion parameters.

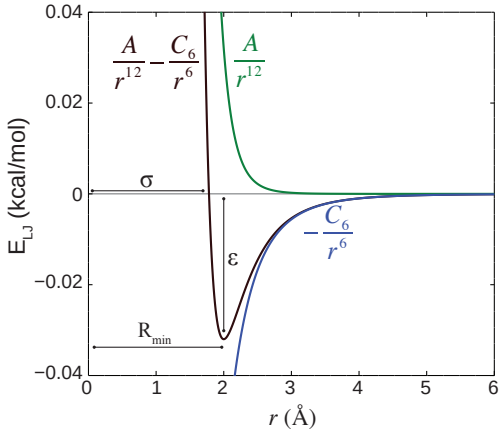

Figure 1: The Lennard-Jones potential for the interaction between two carbon atoms (CGenFF atom type CG311). The total potential is shown in black. The repulsive $1/r^{12}$ component is shown in green and the $1/r^6$ London dispersion component is shown in blue.

Molecular dynamics simulations of organic molecules often employ a generalized molecular-mechanical force field. These force fields define parameters for the standard types of chemical bonds and functional groups present in organic molecules, making it possible to generate a force field automatically for an arbitrary molecule. CGenFF,[27] GAFF,[28] and OPLS[29] are popular generalized force fields. More recently, models that are capable of describing induced polarization, such as Drude polarizable force fields, have also been developed.[30]

The CGenFF, GAFF, OPLS, and Drude force fields all account for inter-atomic London dispersion interactions via the Lennard-Jones 12-6 potential. This potential combines the

attractive $1/r^6$ London dispersion term with a $1/r^{12}$ short-range repulsive term.[31] The components of this potential are illustrated in Figure 1. Higher order dispersion terms (i.e., C_8 and C_{10}) are neglected. The Lennard-Jones potential for the interaction between atoms i and j is

$$E_{LJ,ij}(r_{ij}) = \frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{6,ij}}{r_{ij}^{6}} \tag{5}$$

Appropriate A and C_6 coefficients must be defined for each pair of atoms in the system. This equation is more commonly formulated as

$$E_{LJ,ij}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \tag{6}$$

where ϵ_{ij} is the potential well depth and σ_{ij} is the sum of atomic radii at which the potential crosses the zero of energy.

To reduce the number of parameters needed to define the force field, each atom in the system is assigned an atom type. All atoms of the same type are assumed to have the same Lennard-Jones parameters. The type of an atom is generally specified by its element, hybridization, and bonding partners. The number and definition of atom types varies widely between different force fields. For instance, for the molecules studied in this paper, the CGenFF, GAFF, and OPLS force fields have 67, 32, and 153 atom types, respectively.

The σ and ϵ parameters for molecular-mechanical force fields are typically assigned by performing simulations of bulk liquids using various parameter sets. For each set of parameters, properties like the density and enthalpy of vaporization of the neat liquid are calculated. The parameters that yield the most accurate properties are used as the standard Lennard-Jones potential for that atom type. While this practice has been effective, there are several associated drawbacks. This procedure assumes the Lennard-Jones parameters are transferable to atoms of the same type in other molecules. Additionally, fitting parameters to properties of bulk liquids becomes more difficult for polyatomic molecules because the Lennard-Jones parameters for multiple atom types must be fit simultaneously. Both ϵ

and σ are treated as free parameters of empirical force fields, along with hundreds of other parameters. This creates the possibility that the parameterization procedure will generate unphysical values for ϵ and σ.

The need for greater accuracy in molecular simulations has spurred efforts to validate force field parameters. The Virtual Chemistry database provides structures and topology files for simulations of molecular liquids with the CGenFF, GAFF, and OPLS force fields. This provides an extensive test set to evaluate the accuracy of the force-field parameters. Simulations of molecular liquids in this test set have shown that the computational predictions can be significantly in error for some properties, although it is not always apparent which parameter(s) require adjustment. The ability of XDM to compute atomic dispersion coefficients from first principles provides a novel way of determining if the C_6 dispersion coefficients of a force field are physically realistic.

This paper presents the calculation of C_6 coefficients using the XDM model on 88 molecules from the Virtual Chemistry force-field test set. The calculated coefficients are compared to those derived from the Lennard-Jones parameters for the CGenFF, GAFF, OPLS, and Drude force fields. A revised force field for liquid benzene is derived based on the XDM dispersion coefficients.

2 Computational Methods

2.1 XDM Calculations

XDM dispersion coefficients were calculated for a set of 88 molecules. Gas-phase structures from the Virtual Chemistry database were taken as the initial geometries. These structures were energy minimized with the PBE0 functional[32] and the def2-SVP basis set[33] using ORCA 3.0.[34] Further geometry optimization was then performed with PBE0/aug-cc-pVTZ using Gaussian 09.[35] Single-point energy calculations with PBE0/aug-cc-pVTZ were carried out to generate the wavefunction files needed to determine the XDM dispersion coefficients. This

method has been shown to be reliable for predicting the molecular electrostatic properties of small molecules.[36] The XDM dispersion coefficients were calculated from the PBE0/aug-cc-pVTZ wavefunction files using the postg program.[25,37,38] A Python script that automates the parsing of dispersion coefficients from the output of postg is available through GitHub.[39] Force field parameters for molecules in the Virtual Chemistry test set were extracted from the published itp files.[40] The equations for conversion of these parameters to a C_6 coefficient in atomic units are given in the appendix. Sample input files are included in Supporting Information.

2.2 Molecular Dynamics Simulations

Simulations to parameterize molecular-mechanical force fields were performed using GROMACS 5.1.4.[41] The simulations were performed under periodic boundary conditions with unit cells containing 1000 molecules. Where possible, initial coordinates were taken from the Virtual Chemistry database. In the remaining cases, initial coordinates were generated using the GROMACS insert-molecules module. A Parrinello–Rahman barostat[42,42] and Nosé–Hoover thermostat[43,44] were used in order to sample the isothermal-isobaric ensembles. Electrostatic interactions were calculated using Particle Mesh Ewald (PME) with a grid spacing of 1 Å. Lennard-Jones interactions were calculated using the lattice-sum method.[45,46] Properties were calculated from a 1 ns simulation to equilibrate the system followed by a 1 ns simulation to sample the properties. Sample input files are included in Supporting Information.

3 Results and Discussion

3.1 Molecular Dispersion Coefficients

The molecular C_6 dispersion coefficients for the molecules in the test set were calculated using XDM and the force-field parameters. The correlations between the molecular XDM

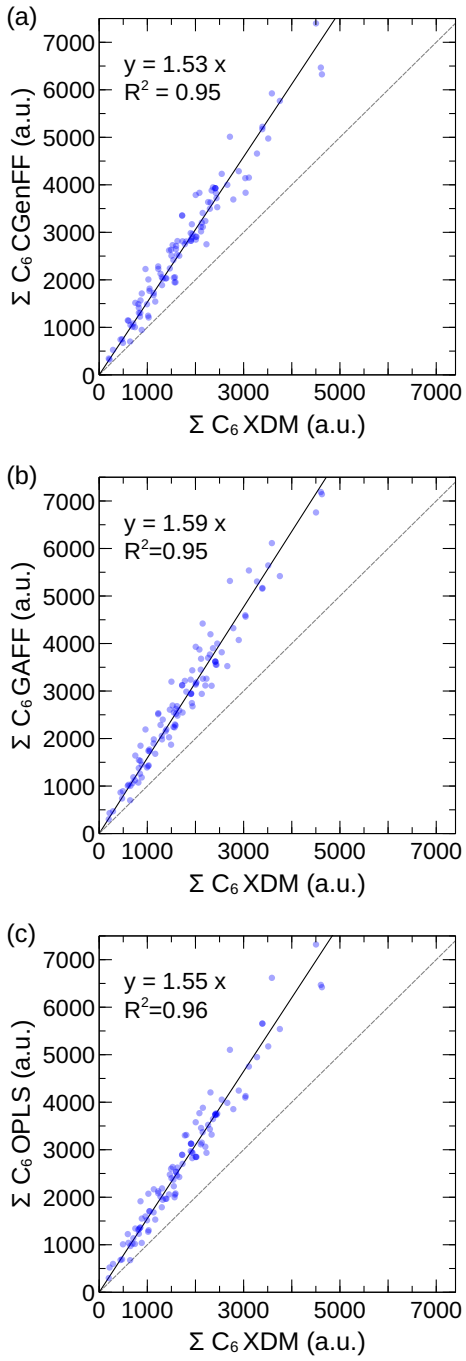

Figure 2: Correlation between the molecular dispersion coefficients of the (a) CGenFF, (b) GAFF, and (c) OPLS force fields and XDM. Each point represents a single molecule of the Virtual Chemistry test set.

dispersion coefficients and the molecular dispersion coefficients for the CGenFF, GAFF, and OPLS force fields are plotted in Figure 2. There is a systematic trend for the force fields to

overestimate the molecular dispersion coefficients, with regression coefficients of 1.53, 1.59, and 1.55 for the CGenFF, GAFF, and OPLS force fields, respectively. This suggests that the dispersion interactions in molecular liquids will be overestimated by these molecular-mechanical force fields.

The overestimation of dispersion coefficients in these force fields may be the result of the neglect of some components of the intermolecular interactions. Because the C_6 coefficients are parameterized to empirical liquid properties, the C_6 coefficients will be assigned spuriously large values to compensate for these neglected intermolecular interactions. For instance, the generalized force-field models use fixed atomic charges to represent electrostatic interactions. This neglects electrostatic interactions due to induced polarization, so the C_6 coefficients of these force fields may have been assigned spuriously large values to compensate for the underestimation of electrostatic interactions. The development of polarizable molecular-mechanical models is one route to address these issues.

These force fields also neglect the 8th- and 10th-order dispersion interactions. Quantum-chemical calculations have shown that the C_8 dispersion term accounts for ca. 30% of the dispersion energy in both molecular dimers[47] and in molecular crystals.[10] The latter corresponds to roughly 20% of the computed lattice energies. A moderate increase in the magnitude of the C_6 dispersion terms can compensate for the neglect of higher-order terms, because these terms are much shorter-range than the 6th-order terms. Non-nonbonded potentials that include higher-order dispersion coefficients have been proposed in the past,[48] but have not been widely adopted. The rigorous inclusion of higher-order dispersion terms in molecular-mechanical force fields may result in more accurate calculation of dispersion interactions and C_6 coefficients that are in better agreement with the experimental and XDM values. XDM provides a first-principles method of obtaining these coefficients, which greatly simplifies the parameterization of these additional terms.

Although many-body effects are sometimes invoked as a neglected source of dispersion energy, analysis using methods like XDM have generally found that pairwise interactions

account for the bulk of the dispersion interaction, while non-additive many-atom dispersion is slightly repulsive in general and only accounts for a very small fraction of the total dispersion energy.[49–51]

Finally, the widespread practice of applying a switching function to terminate the Lennard-Jones interaction at a moderate distance (e.g., 12 Å) has also caused the force field C_6 dispersion coefficients to be exaggerated. This truncation is also used when the force field is parameterized, so the parameterization procedure of some force fields tends to assign a spuriously large C_6 coefficient to the parameterized atoms to compensate for the neglected long-range dispersion interactions (although a correction is made for this approximation in some force fields). Fisher et al. found that the enthalpies of vaporization of liquids in the Virtual Chemistry test set were systematically overestimated when the long-range dispersion interactions were calculated,[52] which is consistent with our conclusion that the dispersion coefficients for these force fields are larger than they should be on physical grounds. For homogeneous systems, there are methods of correcting for the neglect long-range dispersion interactions without explicitly calculating them using a lattice-summation method,[53] although this has not been used universally in force field parameterization. The redevelopment of force fields to include long-range dispersion interactions would likely result in smaller C_6 dispersion parameters.

Table 1: Calculated liquid properties obtained with force fields for which the molecular C_6 coefficients differ from the XDM values by a wide margin. The data are taken from Ref. 54. Several molecules that contain sulfur or bromine atoms have enthalpies of vaporization that are significantly different from the experimental values. Dispersion coefficients are given in atomic units.

Force field	Molecule	$C_{6,\mathrm{FF}}$	$C_{6,\mathrm{XDM}}$	ρ expt.	ρ calc.	ΔH_{vap} expt.	ΔH_{vap} calc.
CGenFF	dibromomethane	2227.7	961.6	2496.8	2435.3	37.67	43.53
GAFF	dibromomethane	2191.2	961.6	2496.8	1962.8	37.67	28.83
GAFF	1-bromobutane	3436.0	1933.3	1275.1	1176.4	36.60	35.98
CGenFF	1,2-ethanedithiol	2630.8	1493.1	1113	1167	41.85	48.16
OPLS	1,2-ethanedithiol	2600.4	1493.1	1113	1157	41.85	46.68

If the force field significantly underestimates or overestimates the molecular C_6 coefficient relative to the XDM value, the simulated physical properties of the liquid are often greatly in error. For example, the GAFF force field overestimates the total dispersion coefficient of compounds containing bromine, such as dibromomethane and 1-bromobutane, by a large margin. As shown in Table 1, the density and enthalpy of vaporization of these compounds are significantly in error when anomalously-large C_6 parameters are used to simulate these liquids. In agreement with our analysis, Adluri et al. showed that the GAFF and CGenFF Lennard-Jones parameters for bromine are not optimal.[55] The bromine dispersion coefficients for the reparameterized force field were closer to the XDM values (206.7 a.u. and 163.1 a.u. for the revised CGenFF and GAFF models, respectively). The CGenFF and OPLS force fields also predict anomalously high molecular dispersion coefficients for 1,2-ethanedithiol. These models overestimate the density and enthalpy of vaporization (Table 1).

3.2 Atomic Dispersion Coefficients

A comparison of the force-field and XDM homoatomic C_6 dispersion coefficients allows the validity of force-field atom typing to be assessed and reveals whether the systematic overestimation of the dispersion coefficients can be traced to particular elements. The average of the homoatomic dispersion coefficients for each element, and the accompanying standard deviation, were calculated for the full test set and are reported in Table 2. The force-field dispersion coefficients for each chemically-unique atom in the test set are plotted against the equivalent XDM dispersion coefficients in Figure 3. A more narrow distribution that is restricted to H, C, N, and O atoms is presented in Figure 4.

In general, the XDM dispersion coefficients span a modest range for a given element, with coefficients of variation that range from 0.01 for fluorine to 0.11 for oxygen and nitrogen. This suggests that, for organic compounds, variations in the chemical environment of an atom do not drastically affect the strength of its dispersion interactions (although this is not the case for metals[15] and for changing oxidation states[23]). Thus, physically-realistic force

fields should only have moderate deviations from the average dispersion coefficient for each element. The comparisons of the force-field and XDM homoatomic dispersion coefficients are discussed for each element throughout this section.

Table 2: Average values of the homoatomic C_6 dispersion coefficients, grouped by element, for the compounds in the Virtual Chemistry test set. The breadth of the distribution of dispersion coefficients for a given element is indicated by its standard deviation. For the XDM values, this represents the spread of dispersion coefficients for an element due to the variety of chemical environments in the test set. For the MM models, this reflects that various atom types with different Lennard-Jones parameters are used to represent the same element.

Element	XDM	CGenFF	GAFF	OPLS
H	2.5 ± 0.2	1.7 ± 0.7	1.2 ± 0.5	1.7 ± 0.7
C	22.0 ± 1.3	40.4 ± 9.1	44.4 ± 5.2	39.4 ± 9.8
N	15.9 ± 1.8	44.3 ± 20.9	58.2 ± 0.0	59.6 ± 6.1
O	12.6 ± 1.4	34.6 ± 16.7	43.7 ± 6.2	38.3 ± 7.2
F	8.2 ± 0.1	9.9 ± 0.8	16.3 ± 0.0	9.9 ± 0.8
S(II)	91.8 ± 2.6	212.7 ± 54.4	149.0 ± 0.0	205.9 ± 55.5
S(IV)	58	208	145	149
Cl	70.8 ± 1.5	154.0 ± 3.8	135.0 ± 0.0	135.0 ± 0
Br	134.2 ± 2.2	238.0 ± 0.0	356.0 ± 0.0	238.0 ± 0

3.2.1 Hydrogen

The force-field dispersion coefficients for hydrogen atoms are systematically underestimated relative to the XDM values. As shown in Table 2, the average XDM C_6 coefficient for hydrogen atoms is 2.5 a.u., which is higher than the averages of 1.7 a.u., 1.2 a.u., and 1.7 a.u. for the CGenFF, GAFF, and OPLS force fields, respectively. As hydrogen atoms have relatively weak dispersive and repulsive interactions in comparison to their parent atoms, Lennard-Jones parameters of hydrogen atoms are sometimes assigned somewhat arbitrary parameters or are even neglected in some cases.

XDM calculates atom-in-molecule dispersion coefficients by using the Hirshfeld[56] method to partition the molecular electron density into atomic regions, although the choice of parti-

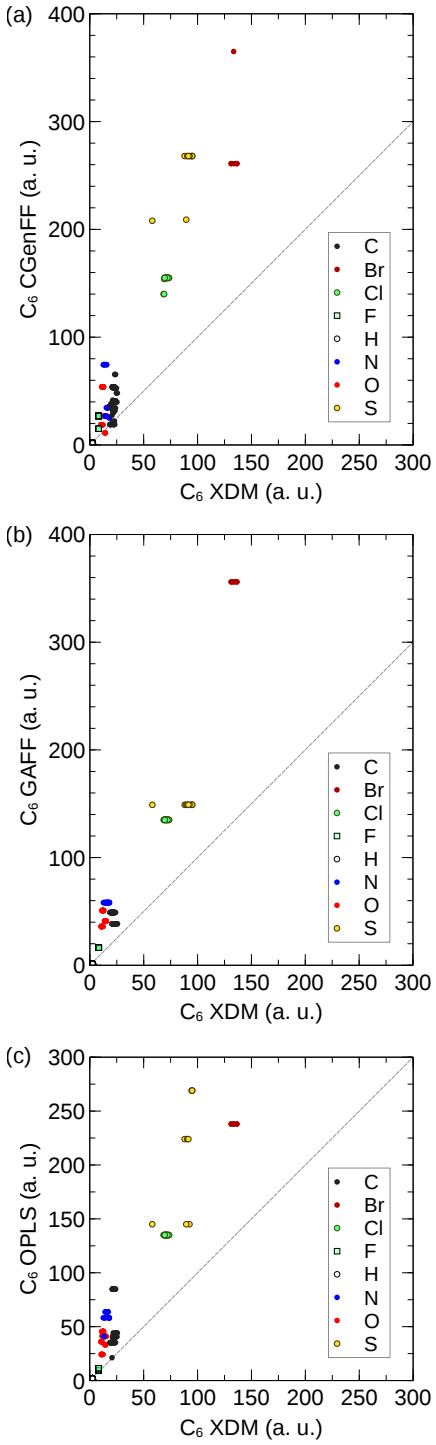

Figure 3: Correlation between the homoatomic dispersion coefficients of the (a) CGenFF, (b) GAFF, and (c) OPLS force fields and XDM. Each point represents a unique atom from the Virtual Chemistry set.

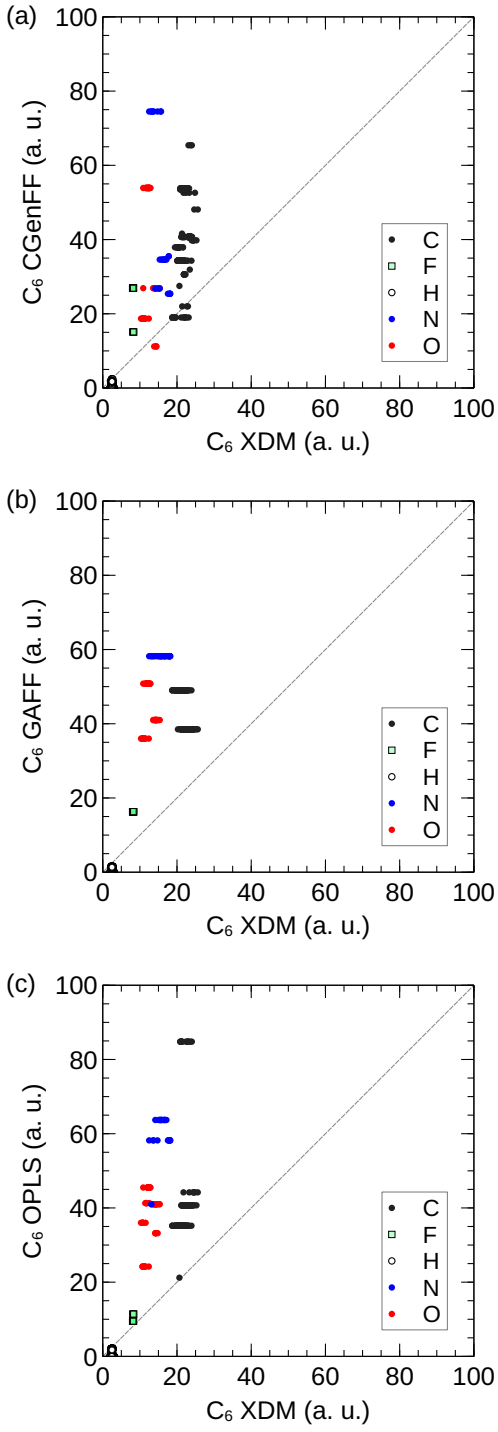

Figure 4: Correlation between the homoatomic dispersion coefficients of the (a) CGenFF, (b) GAFF, and (c) OPLS force fields and XDM, restricted to $C_{6,\mathrm{XDM}} = [0, 100$ a.u.], enclosing H, C, N, O, and F atoms.

tioning scheme is somewhat arbitrary. Partitioning of the electron density between hydrogen atoms and their parent atoms is particularly sensitive to the method used. The choice of the Hirshfeld method causes the hydrogen atoms to account for a greater portion of the molecular dispersion interactions than do the force fields. Newer methods, like the Iterative Hirshfeld Method,[57] could yield hydrogen C_6 dispersion coefficients closer to the force-field values. The net dispersion interaction for an atom and its bound hydrogens is more consistent between different partitioning schemes. These grouped dispersion coefficients also show a systematic overestimation of dispersion coefficients by the force-field models (see Supporting Information).

3.2.2 Carbon

Carbon atoms have sizable dispersion coefficients due to their high polarizability ($\langle C_{6,\mathrm{XDM}} \rangle$ = 22.0±1.3 a.u.). Bonding partners have the largest effect on the dispersion coefficient of carbon atoms. As shown in Figure 5, the XDM dispersion coefficient of the electron-poor tertiary carbon of *t*-butylamine is particularly low, with a value of 18.7 a.u. At the other extreme, the β-carbon of 1,1-dichloroethene has a C_6 coefficient of 25.6 a.u.

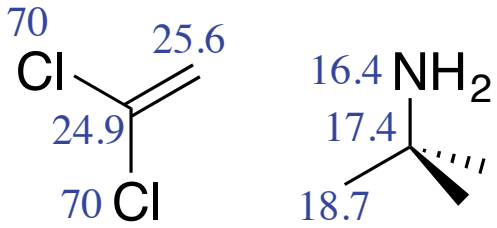

Figure 5: The atomic XDM dispersion coefficients (blue, a.u.) for electron-rich carbon atoms, such as those in 1,1-dichloroethene, and those of electron-poor carbon atoms like *t*-butylamine.

The force fields all overestimate the majority of the carbon dispersion coefficients, with the average values roughly a factor of two higher than the XDM values (Table 2). Moreover, the force fields give a much larger range of dispersion coefficients for carbon than obtained with XDM. The CGenFF force field shows a particularly broad range; the lowest C_6 value

is 19.0 a.u. for methyl-group carbons (atom type CG311) and the highest is 65.4 a.u. for carbonyl carbons (atom type CG2O1). Of the three force fields, GAFF has the most narrow range of C_6 coefficients for carbon, but still gives a standard deviation of $\sigma = 5.2$ a.u. compared to $\sigma = 1.3$ a.u. with XDM.

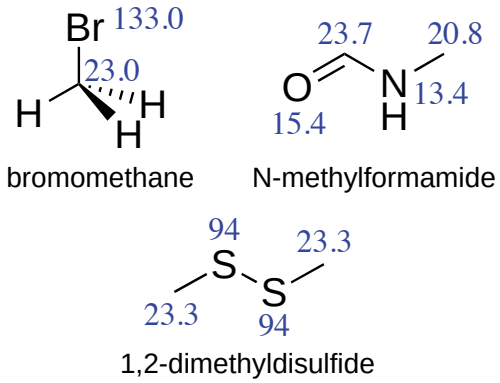

Figure 6: Examples of XDM dispersion coefficients (a.u.) for methyl-group carbon atoms.

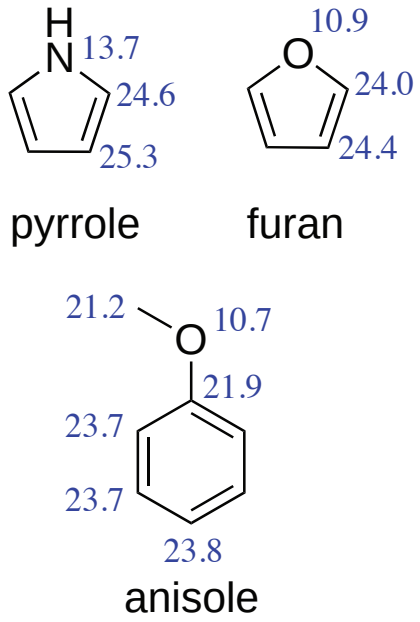

Figure 7: Examples of XDM dispersion coefficients (a.u.) for carbon atoms in aromatic systems.

Conversely, within some force-field carbon atom types, the XDM dispersion coefficients show significant variation. For instance, the CGenFF CG331 atom type represents all methyl

groups. This type includes the electron-deficient methyl group of N-methylformamide, which has an XDM dispersion coefficient of 20.8 a.u. (Figure 6). At the other extreme, the XDM-computed dispersion coefficient is 23 a.u. for carbons bound to bromine or sulfur, as in dimethyl sulfide or bromoethane. The XDM dispersion coefficients also span a significant range for aromatic carbons (Figure 7), such as those represented by the CGenFF CG2R61 atom type. The computed XDM coefficients for this atom type range between 21.5–25.5 a.u. The dispersion coefficients of carbon atoms in electron-rich heteroaromatics, like furan and pyrrole, range between 24–25 a.u., which is incrementally higher than the value of 23.6 a.u. calculated for benzene. There is also a significant variation in the C_6 coefficients of aromatic carbons due to substituent effects. For example, the ipso carbon of anisole has a C_6 coefficient of 21.9 a.u., but the para-carbon has a dispersion coefficient of 23.8 a.u. As force fields use the same dispersion coefficients for all atoms of the same type, the calculated strength of interatomic dispersion interactions could be in error by up to 10% due to the use of atom types.

3.2.3 Nitrogen

XDM predicts nitrogen atoms to have dispersion interactions of moderate strength, with an average C_6 coefficient of 15.9 a.u. Nitrogen atoms have one of the broadest distributions of C_6 coefficients for this test set; electron-poor amide nitrogens have particularly low dispersion coefficients (e.g. 13.4 a.u. in N-methylformamide), while electron-rich alkyl nitrogens have high dispersion coefficients (e.g. 17.2 a.u. in ethane-1,2-diamine). The XDM dispersion coefficients for these molecules are presented in Figure 8.

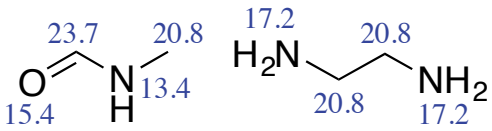

Figure 8: Examples of XDM dispersion coefficients (a.u.) for an electron-poor amide nitrogen (N-methylformamide) and an electron-rich amine nitrogen (ethane-1,2-diamine).

The dispersion coefficients for nitrogen atoms are systematically overestimated by the GAFF and OPLS force fields, which give average C_6 values of 58.2 a.u. and 59.6 a.u., respectively. The CGenFF coefficients for amines are somewhat closer to the XDM values (12–34 a.u for CGenFF vs 14–17 a.u. for XDM), but the coefficients for amide nitrogens (i.e., NG2S0, NG2S1, NG2S1, NG2S2, and NG2S3 atom types) are assigned an anomalously large value (74.5 a.u.). The XDM coefficients for these atoms are actually lower than the average for nitrogen atoms, ranging from 12–15 a.u. The CGenFF amide-nitrogen parameters are shared with the protein backbone of the CHARMM36 force field.[58] The Lennard-Jones parameters for these amide nitrogens were adjusted to provide more accurate backbone hydrogen bonding, but this appears to have caused the C_6 dispersion coefficient to be anomalously high. Although this modification of the Lennard-Jones parameters may describe the energetics of short-range interactions more accurately, the long-range dispersion interactions will be overestimated as a result.

3.2.4 Oxygen

The XDM C_6 coefficients for oxygen are generally smaller than those of carbon and nitrogen ($\langle C_{6,O} \rangle$ = 12.6 a.u.) and span a fairly narrow range of values (σ_O = 1.4 a.u.). Electron-rich carbonyl oxygens have coefficients in the 13–15 a.u range, while alcohols have smaller coefficients, between 12–12.5 a.u. The oxygen C_6 parameters from the three force fields are considerably larger than the XDM values. CGenFF has the smallest oxygen dispersion coefficients, with an average of 34.6 a.u. Carbonyl and ether oxygens are only moderately overestimated ($C_6 < 30$ a.u.), but the coefficients for alcohols are 4–5 times larger than the XDM values (e.g., 53.9 a.u. for OG311). The average C_6 coefficients of the GAFF and OPLS force fields are systematically higher than the XDM values across most oxygen atom types, by factors of 3.5 and 3, respectively. For these force fields, the strengths of dispersion interactions for oxygen atoms are similar to those for carbon atoms, despite oxygen's much lower polarizability.

3.2.5 Sulfur

For sulfur, results for the II and IV oxidation states will be considered separately due to the large effect of oxidation state on C_6 dispersion coefficients.[23] The average XDM dispersion coefficient for S(II) atoms is 91.8 a.u. and the standard deviation is 2.6 a.u. The coefficient averages for CGenFF and OPLS models are overestimated by more than a factor of two and the range of values are extremely large, with standard deviations of 54.5 and 55.5 a.u., respectively. For the CGenFF force field, this is because of the difference between the disulfide and sulfide atom types; the SG301 atom type (C–S–S–C) has a dispersion coefficient of 209 a.u, while the SG311 atom type (SH, –S–) has a dispersion coefficient of 268 a.u.

Sulfolane is the only compound in the test set where the sulfur atom is in the IV oxidation state. The XDM C_6 coefficient for this atom (58 a.u.) is significantly lower than for the S(II) atoms. The CGenFF and OPLS force fields assign the S(IV) atom type (SG302) a dispersion coefficient of 208 a.u., which is lower than for S(II), although it is still overestimated relative to XDM. The GAFF force field assigns S(IV) the same Lennard-Jones parameters as S(II). Given that the sulfur dispersion coefficients are very sensitive to the oxidation state, several distinct sets of Lennard-Jones parameters should be determined for this element.

3.2.6 Halogens

Fluorine has a small XDM dispersion coefficient ($\langle C_{6,F} \rangle$ = 8.2 a.u.), consistent with its low atomic polarizability. The standard deviation of these values is small (σ = 0.1 a.u.), suggesting that a single set of Lennard-Jones parameters are appropriate for fluorine atoms in organofluorines. The fluorine CGenFF and OPLS force-field dispersion coefficients are in reasonable agreement with the XDM values, although the dispersion coefficients for GAFF are roughly double the XDM value.

Chlorine has a large average XDM dispersion coefficient ($\langle C_6 \rangle$ = 70.8 a.u.), but also has a relatively small standard deviation (σ = 1.0 a.u.). The force-field dispersion coefficients for chlorine are much larger than the XDM values. The GAFF and OPLS force field share

Lennard-Jones parameters for chlorine, so their C_6 coefficients are also the same ($C_6 = 135$ a.u.), while the CGenFF force field assigns even larger values ($\langle C_{6,Cl} \rangle = 154$ a.u.).

Bromine has the largest XDM dispersion coefficients in the test set ($\langle C_6 \rangle = 134$ a.u.). The bromine dispersion coefficients are overestimated by all three force fields. In particular, the GAFF force field grossly overestimates the strength of the bromine C_6, with an average value of 356 a.u. The performance of GAFF for the physical properties of organobromine liquids is notably poor, suggesting that the GAFF Lennard-Jones parameters are not physically realistic. Adluri et al. showed that the GAFF model for bromomethane could be improved by reparameterizing the Lennard-Jones terms for Br.[55]

3.3 The Drude Force Field

The Drude force field incorporates the effect of induced polarization by adding charged "Drude" particles that are harmonically tethered to their parent atoms.[59] This model uses a Lennard-Jones potential to represent dispersion interactions, although the parameters for these models generally need to be refit so that they are appropriate for molecules interacting through different Coulombic-interaction terms that are present in polarizable force fields. To evaluate the dispersion parameters of the Drude model, the XDM dispersion coefficients were calculated for 73 molecules from the July 2015 revision of the Drude force field. The correlation between the Drude and XDM C_6 coefficients is presented in Figure 9.

The correlation between the XDM and Drude molecular C_6 coefficients is poorer than for the non-polarizable models; the regression coefficient is 1.67 and the coefficient of determination is 0.81 (Figure 9 (b)). The distribution of atomic dispersion coefficients for the Drude force field is extremely wide, with values ranging from 20 a.u. to more than 100 a.u. for some C, N, and O atom types (Figure 9 (a)). The dispersion coefficients for sulfur, spanning between 200 and 300 a.u., are also overestimated by a large margin. As with the non-polarizable force fields, the systematic overestimation of the dispersion coefficients is likely due to neglect of higher-order C_8 terms and premature truncation of the dispersion-energy

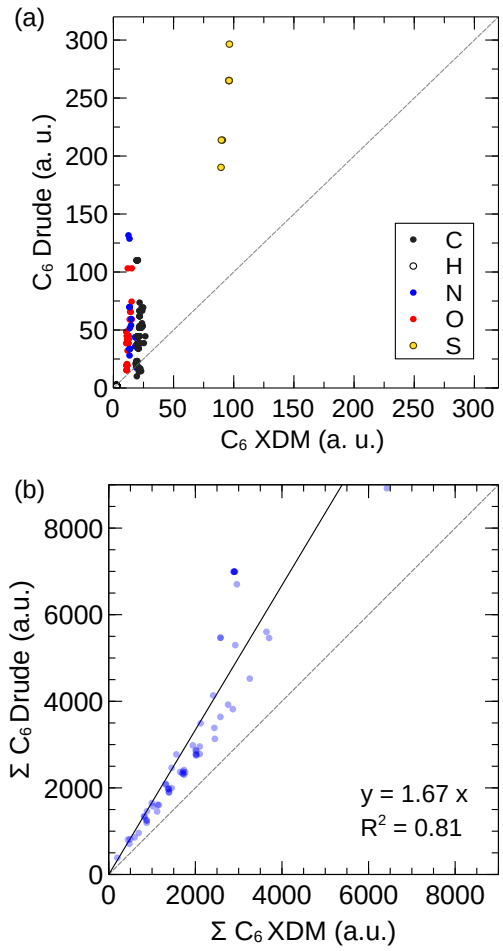

Figure 9: (a) Correlation between the homoatomic dispersion coefficients calculated using the Drude force field and XDM. Each point represents a unique atom from the test set of molecules available in the current version of the Drude force field. (b) Correlation between the Drude and XDM molecular C_6 dispersion coefficients. Each point represents a single molecule of the test set.

summation.

The poor correlation between Drude and XDM dispersion coefficients is exacerbated by five atom types where the force-field dispersion coefficient is several times larger than the corresponding XDM value (Table 3). Atypically large Lennard-Jones ϵ parameters have been assigned to these atoms. For example, atom type ND2R5D, which represents the nitrogen at the 9 position of purines, has an ϵ parameter of -0.23 kcal/mol, which is a factor of 2 larger than is typical for nitrogen atom types. The procedure that determined these parameters included unconventional target data, such as QM interaction energies and

lattice energies.[60–62] These terms are less sensitive to the Lennard-Jones parameters than traditional parameterization target data, like the density or enthalpy of vaporization, so it is possible that the parameters were over-fit. Imposing constraints on the Lennard-Jones parameters to ensure that the C_6 coefficients are in the typical range for each element would be a simple way of avoiding parameters that cause the long-range dispersion interactions to become unrealistically large.

Table 3: Homoatomic dispersion coefficients (in atomic units) for Drude atom types where the force-field C_6 is much larger than the corresponding XDM value.

Atom type	Environment	XDM	Drude
CD31FA	C1 carbon of carbohydrates	19.8	110.1
ND2R5D	GUA/ADE 5-member ring	13.4	128.7
OD2C2B	carboxylate O, anionic phosphate, lipids	12.0	103.2
ND2A3	amide, tertiary DMA	12.5	131.7
SD31A	alkyl thiol sulfur	96.4	298.5

3.4 Water Models

One of the most common applications of generalized force fields is to simulate organic solutes in an aqueous solution. A wide range of water models are available, including models with charges at the three atomic centers (e.g., TIP3P), models with an off-center charge (e.g., TIP4P), and polarizable water models (e.g., SWM4-NDP). The generalized force fields evaluated here are typically used with the TIP3P water model. The TIP3P model overestimates the dielectric constant and diffusivity of water,[63] so there has been interest in adopting water models that describe the properties of water more accurately. To describe solvation properly, the dispersion coefficients of the water model must be balanced with those of the solute force field. The C_6 dispersion coefficients for 11 popular water models are presented in Table 4.

A molecular C_6 coefficient for H_2O of 45.4 a.u. has been estimated by Zeiss and Meath using photoabsorption and high energy inelastic scattering experiments.[64] XDM is in close agreement with this value, with a predicted C_6 coefficient of 45.8 a.u. The TIP3P, TIP3P-FB,

SPCE, OPC3, TIP4P, and TIP4P-Ew water models have C_6 coefficients that are close to the experimental value, ranging from 43 to 49 a.u. The TIP4P-FB, TIP4P/2005, and SWM6-NDP models modestly overestimate the dispersion coefficient, ranging from 50–55 a.u. The SWM4-NDP and TIP4P-D models overestimate the C_6 coefficient by a large margin, with values of 63.7 and 65.3 a.u., respectively.

Our analysis of the molecular C_6 coefficients of the generalized force fields in Section 3.1 indicated that the dispersion interactions are systematically overestimated in comparison to the predictions by XDM. In contrast, many of the water models used in combination with these force fields have dispersion coefficients that are comparable to the water dispersion coefficient calculated by XDM or determined experimentally. This suggests that the coefficients for the dispersion interaction between TIP3P-model water and a solute described using one of these generalized force field could be unbalanced. Best et al. found that water–protein interactions were predicted to be too strong and had to be attenuated to describe intrinsically disordered proteins correctly.[65]

Table 4: Molecular C_6 coefficients for various molecular-mechanical water models that include a Lennard-Jones 12-6 potential to represent dispersion. The XDM value was calculated for a single gas-phase water molecule using PBE0/aug-cc-pVTZ.

Water model	C_6 (a.u.)	Ref.
TIP3P	43.2	66
TIP3P-FB	46.6	67
SPCE	45.4	68
TIP4P	44.3	66
TIP4P-Ew	47.4	69
TIP4P-FB	52.3	67
TIP4P/2005	53.4	70
TIP4P-D	65.3	71
OPC3	48.5	72
SWM4-NDP	63.7	73
SWM6	50.3	74
XDM	45.8	this work
exptl.	45.4	64

3.5 Force Field Development using XDM-Derived Dispersion Coefficients

To test if XDM-derived C_6 coefficients can be used to parameterize a molecular-mechanical force field, we developed new Lennard-Jones parameters for benzene, toluene, cyclohexane, n-pentane, and n-hexane using XDM data. The internal energy terms and atomic charges of the CGenFF force field models of these molecules were used without modification, but the atomic Lennard-Jones parameters were selected such that the molecular C_6 dispersion coefficient is near the XDM value.

To perform the fitting for benzene, toluene, and cyclohexane, a 4-dimensional grid of ε_H, σ_H, ε_C, and σ_C was considered for each unique C–H pair of CGenFF atom types (see Table 6). The LJ parameters for the sp^2 carbons and bonded hydrogens for benzene were transferrable to toluene and, as such, only the additional parameters for the methyl-group atoms were fit to the toluene reference data. Parameter sets yielding a molecular dispersion coefficient that deviated from the XDM value by more than a given threshold were discarded (see Table 6). Molecular dynamics simulations were performed using the remaining parameter sets to calculate the density and enthalpy of vaporization of the liquid. Dispersion interactions were calculated using the LJ-PME method so that long-range dispersion interactions were included.[52] The enthalpy of vaporization was calculated from the average potential energy the liquid simulation,

$$\Delta H_{vap} = RT + \langle \mathcal{V}_{pot} \rangle_{gas} - \langle \mathcal{V}_{pot} \rangle_{liquid}. \tag{7}$$

The final, optimum parameters were those that yielded the lowest deviation from experiment based on the target function,

$$\delta(\sigma, \varepsilon) = \left(\frac{\rho_{exptl.} - \rho_{calc}}{\rho_{exptl.}} \right)^2 + \left(\frac{\Delta H_{vap,exptl.} - \Delta H_{vap,calc}}{\Delta H_{vap,exptl.}} \right)^2, \tag{8}$$

The parameters for the linear alkanes were determined by the same XDM-constrained

parameterization procedure, but only the top 5 parameter pairs from the search for the analogous sp^3 carbon atom types from cyclohexane and toluene were considered. This procedure successfully identified effective parameters from only 25 MD simulations of trial parameters. The original and new parameters for all molecules are included in Supporting Information.

Both the density and enthalpy of vaporizations predicted using the new parameters are in good agreement with the experimental values and are in better agreement than the original CGenFF parameters in some cases (Table 5). This is particularly true for the enthalpy of vaporizations. Between 95–98% of potential parameters combinations were excluded by the XDM criteria, making this approach much more efficient than a traditional grid search of the Lennard-Jones parameter space. This suggests that XDM molecular dispersion coefficients can provide bounds for Lennard-Jones parameters that limit the parameter space to physically-realistic values.

Table 5: Physical properties of 5 different liquids calculated using the original CGenFF force field and a reparameterized force field (FF-XDM). Units are kg/m^3 and kcal/mol for density (ρ) and enthalpy of vaporization (ΔH_{vap}), respectively.

Compound	FF-XDM		CGenFF		Exptl.	
	ρ	ΔH_{vap}	ρ	ΔH_{vap}	ρ	ΔH_{vap}
benzene	869	33.6	861	28.5	876	33.9
toluene	872	38.3	864	46.7	865	38.1
cyclohexane	766	31.8	783	35.3	778	32.0
n-pentane	621	20.6	628	11.9	626	26.2
n-hexane	667	34.05	666	27.91	655	31.51

4 Conclusions

The XDM method was used to calculate the C_6 dispersion coefficients of 88 molecules from the Virtual Chemistry test set. These density-functional-theory-derived dispersion coefficients were compared to dispersion coefficients defined through the Lennard-Jones potential parameters of the CGenFF, GAFF, and OPLS molecular-mechanical force fields.

Table 6: Parameterization space for Lennard-Jones parameters of dispersion-bound liquids. The range of Lennard-Jones σ and ϵ parameters in the search for each atom type is given. 10 parameters in that range were evaluated, yielding 10000 potential parameter combinations. The percentage of parameter combinations eliminated by the XDM criteria and the allowable deviation of the force field molecular C_6 from the XDM value are also given.

Compound	Atom type	σ range (Å)	ϵ range (kcal/mol)	Parameter space reduction (%)	XDM deviation threshold
benzene	CG2R61	$[3.380, 3.720]$	$[0.214, 0.372]$	94.86%	5%
	HGR61	$[2.250, 2.590]$	$[0.047, 0.204]$		
toluene	CG331	$[3.553, 3.753]$	$[0.279, 0.374]$	98.69%	1.5%
	HGA3	$[2.288, 2.488]$	$[0.053, 0.148]$		
cyclohexane	CG321	$[3.481, 3.681]$	$[0.187, 0.282]$	98.11%	1%
	HGA2	$[2.288, 2.488]$	$[0.099, 0.194]$		

All three force fields systematically overestimate the molecular dispersion coefficients relative to XDM. The empirical parameterization process likely led to anomalously high C_6 dispersion coefficients due to the neglect of long-range dispersion interactions, higher-order dispersion terms, and induced electronic polarization. Next-generation force fields that account for these terms may return more realistic C_6 dispersion interactions. Improved model for five organic liquids were successfully developed, demonstrating that it is possible to define accurate force fields with physically-realistic C_6 dispersion coefficients while using the standard form of the force field. This procedure can also dramatically reduce the cost of the parameterization by reducing the number of putative parameters to the small subset that are consistent with the XDM-derived dispersion coefficients.

Molecular-mechanical force fields use the same dispersion interaction parameters for all atoms of the same type. In some cases, the XDM dispersion coefficients spanned a significant range of values for atoms of the same type, indicating that the Lennard-Jones parameters are not optimal for some atoms. This is particularly true for the GAFF and OPLS force fields, which have fewer variants of Lennard-Jones parameters than the CGenFF force field. One example where molecular-mechanical atom typing breaks down are methyl carbon atoms in electron-rich versus electron-poor environments, which XDM predicts to have significantly

different dispersion coefficients. The introduction of new atom types motivated by groupings of C_6 values may improve the accuracy of force-field models.

The Drude polarizable force field displayed an even poorer correspondence with the XDM C_6 coefficients than the three non-polarizable models. Several atom types in the Drude force field have dispersion coefficients that are many times greater than the XDM values. It is possible that the increased number of terms in this force field, due to the incorporation of the polarizability, cause the parameters to be underdetermined. The parameter-fitting process could benefit from additional constraints, such as physically-reasonable molecular dispersion coefficients derived from XDM.

Some standard water models, such as TIP3P, have dispersion coefficients that are very similar to the XDM and experimental values. Other models, such as TIP4P-D and SWM4-NDP, overestimate the magnitude of the dispersion coefficient by up to 50%. Thus, the C_6 coefficient may be a worthwhile term to consider in the evaluation of water models in order to ensure the dispersion interactions are physically realistic.

The use of quantum-chemical methods like XDM in the parameterization of dispersion interactions could provide new opportunities to develop more realistic force fields, as illustrated here in the cases of benzene, toluene, cyclohexane, n-pentane, and n-hexane. An immediate application of XDM will be to validate Lennard-Jones parameters of force fields to ensure that the molecular and atomic C_6 dispersion coefficients do not deviate from the XDM values by a large margin. XDM also provides an effective means to calculate the C_8 and C_{10} dispersion coefficients, which are currently neglected from conventional molecular-mechanical force fields. Parameterizing these terms had been impractical because the model is already underdetermined, but XDM could provide reasonable ab initio values. This suggests a general strategy for parameterizing the non-bonded parameters for force fields, where XDM is used to assign the atomic dispersion coefficients. The repulsive component could be derived from a QM potential energy surface, AIMD simulation, topological analysis of the electron density, or empirical fitting from molecular dynamics simulations of condensed

states.

5 Appendix A: Conversion of Dispersion Coefficients

The postg XDM code directly reports C_6 dispersion coefficients in atomic units. Parameter files for GROMACS and CHARMM store the dispersion coefficients through the parameters for the Lennard-Jones potential (Figure 1). In GROMACS, the LJ potential is defined as

$$E_{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right]. \tag{9}$$

The C_6 coefficient, in terms of these σ and ε parameters, is

$$C_6 = 4\varepsilon\sigma^6. \tag{10}$$

These LJ parameters are given in terms of kJ/mol for ε and nm for σ. The conversion to atomic units is

$$1(\text{kJ/mol})\text{nm}^6 = 17344.659 \text{ a.u.} \tag{11}$$

CHARMM defines the Lennard-Jones potential in terms of the location of the potential-energy minimum, R_{min}, instead of σ.

$$E_{LJ}(r) = \varepsilon \left[\left(\frac{R_{min}}{r}\right)^{12} - 2\left(\frac{R_{min}}{r}\right)^{6} \right] \tag{12}$$

$$= \varepsilon\frac{R_{min}^{12}}{r^{12}} - 2\varepsilon\frac{R_{min}^{6}}{r^{6}}, \tag{13}$$

In terms of R_{min}, C_6 is defined as

$$C_6 = 2\varepsilon {R_{min}}^6. \tag{14}$$

These LJ parameters are given in terms of kcal/mol for ε and Å for R_{min}.[1] The conversion

[1] CHARMM-format parameter files actually store $R_{min}/2$

to atomic units is

$$1(\text{kcal/mol})\text{Å}^6 = 0.07257 \text{ a.u.} \tag{15}$$

Acknowledgement

The authors thank NSERC of Canada for funding through the Discovery Grants program (Application 418505-2012 and RGPIN-05795-2016). MM thanks the School of Graduate Studies at Memorial University for a graduate fellowship. Computational resources were provided by Compute Canada (RAPI: djk-615-ab) through the Calcul Quebec, WestGrid, and ACENET consortia. We thank Archita Adluri and Evan Walters for proofreading the manuscript. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

Supporting Information Available

CSV file of average C_6, C_8, and C_{10} coefficients for each atom type, sample input files for ORCA, Gaussian 09, and GROMACS. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) London, F. On the theory and systematic of molecular forces. *Z. Physik* **1930**, *63*, 245–279.

(2) Missner, A.; Kügler, P.; Antonenko, Y. N.; Pohl, P. Passive transport across bilayer lipid membranes: Overton continues to rule. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, E123.

(3) Missner, A.; Pohl, P. 110 Years of the Meyer–Overton Rule: Predicting Membrane

Permeability of Gases and Other Small Compounds. *ChemPhysChem* **2009**, *10*, 1405–1414.

(4) Riahi, S.; Rowley, C. N. Why Can Hydrogen Sulfide Permeate Cell Membranes? *J. Am. Chem. Soc.* **2014**, *136*, 15111–15113.

(5) Awoonor-Williams, E.; Rowley, C. N. Molecular simulation of nonfacilitated membrane permeation. *Biochim. Biophys. Acta – Biomembranes* **2016**, *1858*, 1672–1687.

(6) Zhou, R.; Huang, X.; Margulis, C. J.; Berne, B. J. Hydrophobic Collapse in Multidomain Protein Folding. *Science* **2004**, *305*, 1605–1609.

(7) Barratt, E.; Bingham, R. J.; Warner, D. J.; Laughton, C. A.; Phillips, S. E. V.; Homans, S. W. Van der Waals Interactions Dominate Ligand–Protein Association in a Protein Binding Site Occluded from Solvent Water. *J. Am. Chem. Soc* **2005**, *127*, 11827–11834.

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

(9) Kolář, M.; Kubař, T.; Hobza, P. On the Role of London Dispersion Forces in Biomolecular Structure Determination. *J. Phys. Chem. B.* **2011**, *115*, 8038–8046.

(10) Otero-de-la Roza, A.; Johnson, E. R. A benchmark for non-covalent interactions in solids. *J. Chem. Phys.* **2012**, *137*, 054103.

(11) Reilly, A. M.; Tkatchenko, A. Understanding the role of vibrations, exact exchange, and many-body van der Waals interactions in the cohesive properties of molecular crystals. *J. Chem. Phys.* **2013**, *139*, 024705.

(12) Whittleton, S. R.; Otero-de-la-Roza, A.; Johnson, E. R. The exchange-hole dipole dispersion model for accurate energy ranking in molecular crystal structure prediction. *J. Chem. Theory Comput.* **2017**, *13*, 441–450.

(13) Beran, G. J. O. Modeling Polymorphic Molecular Crystals with Electronic Structure Theory. *Chem. Rev.* **2016**, *116*, 5567–5613.

(14) Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. A Major Advance in Crystal Structure Prediction. *Angew. Chem. Int. Ed.* **2008**, *47*, 2427–2430.

(15) Christian, M. S.; Otero-de-la-Roza, A.; Johnson, E. R. Surface adsorption from the exchange-hole dipole moment dispersion model. *J. Chem. Theory. Comput.* **2016**, *12*, 3305–3315.

(16) Christian, M. S.; Otero-de-la-Roza, A.; Johnson, E. R. Adsorption of graphene to nickel (111) using the exchange-hole dipole moment model. *Carbon* **2017**, *118*, 184âĂŞ191.

(17) Stone, A. *The Theory of Intermolecular Forces*, 2nd ed.; Oxford University Press: Oxford, 2013.

(18) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.

(19) Cole, D. J.; Vilseck, J. Z.; Tirado-Rives, J.; Payne, M. C.; Jorgensen, W. L. Biomolecular Force Field Parameterization via Atoms-in-Molecule Electron Density Partitioning. *J. Chem. Theory Comput.* **2016**, *12*, 2312–2323.

(20) Vandenbrande, S.; Waroquier, M.; Speybroeck, V. V.; Verstraelen, T. The Monomer Electron Density Force Field (MEDFF): A Physically Inspired Model for Noncovalent Interactions. *J. Chem. Theory Comput.* **2017**, *13*, 161–179.

(21) Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction revisited. *J. Chem. Phys.* **2007**, *127*, 154108.

(22) Johnson, E. R. In *Non-covalent Interactions in Quantum Chemistry and Physics*; Otero-de-la Roza, A., DiLabio, G. A., Eds.; Elsevier, 2017; Chapter 5, pp 215–248.

(23) Johnson, E. R. Dependence of dispersion coefficients on atomic environment. *J. Chem. Phys.* **2011**, *135*, 234109.

(24) Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction: High-order dispersion coefficients. *J. Chem. Phys.* **2006**, *124*, 014104.

(25) Otero-de-la Roza, A.; Johnson, E. R. Non-Covalent Interactions and Thermochemistry using XDM-Corrected Hybrid and Range-Separated Hybrid Density Functionals. *J. Chem. Phys.* **2013**, *138*, 204109.

(26) Otero-de-la Roza, A.; Johnson, E. R. Van der Waals interactions in solids using the exchange-hole dipole moment. *J. Chem. Phys.* **2012**, *136*, 174109.

(27) 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.

(28) 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.

(29) 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.

(30) Lopes, P.; Roux, B.; MacKerell Jr., A. D. Molecular Modeling and Dynamics Studies with Explicit Inclusion of Electronic Polarizability: Theory and Applications. *Theor. Chem. Acc.* **2009**, *124*, 11–28.

(31) Jones, J. E. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proc. R. Soc. London, A* **1924**, *106*, 463–477.

(32) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(33) 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.

(34) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.

(35) 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.

(36) Hickey, A. L.; Rowley, C. N. Benchmarking Quantum Chemical Methods for the Calculation of Molecular Dipole Moments and Polarizabilities. *J. Phys. Chem. A* **2014**, *118*, 3678–3687.

(37) Kannemann, F. O.; Becke, A. D. Atomic volumes and polarizabilities in density-functional theory. *J. Chem. Phys.* **2012**, *136*, 034109.

(38) The postg program is available from http://schooner.chem.dal.ca.

(39) PyPostg. `https://github.com/RowleyGroup/postg-parser`, (accessed March 17, 2017).

(40) Virtual Chemistry. `http://virtualchemistry.org/`, (accessed March 17, 2017).

(41) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

(42) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, *45*, 1196–1199.

(43) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511–519.

(44) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.

(45) Wennberg, C. L.; Murtola, T.; Hess, B.; Lindahl, E. Lennard-Jones Lattice Summation in Bilayer Simulations Has Critical Effects on Surface Tension and Lipid Properties. *J. Chem. Theory Comput.* **2013**, *9*, 3527–3537.

(46) Wennberg, C. L.; Murtola, T.; Páll, S.; Abraham, M. J.; Hess, B.; Lindahl, E. Direct-Space Corrections Enable Fast and Accurate Lorentz–Berthelot Combination Rule Lennard-Jones Lattice Summation. *J. Chem. Theory Comput.* **2015**, *11*, 5737–5746.

(47) Johnson, E. R.; Becke, A. D. A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. *J. Chem. Phys.* **2006**, *124*, 174104.

(48) Pastor, R. W. Topics in stochastic dynamics of polymers. Ph.D. thesis, Harvard University, 1984.

(49) Otero-de-la Roza, A.; Johnson, E. R. Many-body dispersion interactions from the exchange-hole dipole moment model. *J. Chem. Phys.* **2013**, *138*, 054103.

(50) Kennedy, M. R.; McDonal, A. R.; DePrince III, A. E.; Marshall, M. S.; Podeszwa, R.; Sherril, C. D. Resolving the three-body contribution to the lattice energy of crystalline benzene: Benchmark results from coupled-cluster theory. *J. Chem. Phys.* **2014**, *140*, 121104.

(51) Huang, Y.; Beran, G. J. Reliable prediction of three-body intermolecular interactions using dispersion-corrected second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2015**, *143*, 044113.

(52) Fischer, N. M.; van Maaren, P. J.; Ditz, J. C.; Yildirim, A.; van der Spoel, D. Properties of Organic Liquids when Simulated with Long-Range Lennard-Jones Interactions. *J. Chem. Theory Comput.* **2015**, *11*, 2938–2944.

(53) Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Pande, V. S. Accurate and Efficient Corrections for Missing Dispersion Interactions in Molecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 13052–13063.

(54) van der Spoel, D.; van Maaren, P. J.; Caleman, C. GROMACS molecule and liquid database. *Bioinformatics* **2012**, *28*, 752.

(55) Adluri, A. N. S.; Murphy, J. N.; Tozer, T.; Rowley, C. N. Polarizable Force Field with a sigma-Hole for Liquid and Aqueous Bromomethane. *J. Phys. Chem. B* **2015**, *119*, 13422–13432.

(56) Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta* **1977**, *44*, 129–138.

(57) Bultinck, P.; Alsenoy, C. V.; Ayers, P. W.; Carbó-Dorca, R. Critical analysis and extension of the Hirshfeld atoms in molecules. *J. Chem. Phys.* **2007**, *126*, 144111.

(58) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; Alexander D. MacKerell, J. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain $\chi 1$ and $\chi 2$ Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(59) Lamoureux, G.; Roux, B. Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

(60) Zhu, X.; MacKerell Jr., A. D. Polarizable Empirical Force Field for Sulfur-Containing Compounds Based on the Classical Drude Oscillator Model. *J. Comput. Chem.* **2010**, *31*, 2330–2341.

(61) Baker, C. M.; Anisimov, V. M.; MacKerell Jr., A. D. Development of CHARMM Polarizable Force Field for Nucleic Acid Bases Based on the Classical Drude Oscillator Model. *J. Phys. Chem. B* **2011**, *115*, 580–596.

(62) Patel, D. S.; He, X.; MacKerell Jr., A. D. Polarizable Empirical Force Field for Hexopyranose Monosaccharides Based on the Classical Drude Oscillator. *J. Phys. Chem. B* **2015**, *119*, 637–652.

(63) Braun, D.; Boresch, S.; Steinhauser, O. Transport and dielectric properties of water and the influence of coarse-graining: Comparing BMW, SPC/E, and TIP3P models. *J. Chem. Phys.* **2014**, *140*, 064107.

(64) Zeiss, G.; Meath, W. J. The H2O-H2O dispersion energy constant and the dispersion of the specific refractivity of dilute water vapour. *Mol. Phys.* **1975**, *30*, 161–169.

(65) Best, R. B.; Zheng, W.; Mittal, J. Balanced ProteinâĂŞWater Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. *J. Chem. Theory Comput.* **2014**, *10*, 5113–5124.

(66) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(67) 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.

(68) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

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

(70) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(71) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. *J. Phys. Chem. B* **2015**, *119*, 5113–5123.

(72) Izadi, S.; Onufriev, A. V. Accuracy limit of rigid 3-point water models. *The Journal of Chemical Physics* **2016**, *145*, 074501.

(73) Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell Jr., A. D. A Polarizable Model of Water for Molecular Dynamics Simulations of Biomolecules. *Chem. Phys. Lett.* **200**, *418*, 245–249.

(74) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell Jr., A. D. Six-site polarizable model of water based on the classical Drude oscillator. *J. Chem. Phys.* **2013**, *138*, 034508.

Graphical TOC Entry