# Deriving Force-Field Parameters from First Principles Using a Polarizable and Higher Order Dispersion Model

Koen M. Visscher and Daan P. Geerke*

AIMMS Division of Molecular Toxicology, Department of Chemistry and Pharmaceutical Sciences, Faculty of Science, Vrije Universiteit Amsterdam, De Boelelaan 1108, 1081 HZ Amsterdam, the Netherlands

S *Supporting Information*

ABSTRACT: In this work we propose a strategy based on quantum mechanical (QM) calculations to parametrize a polarizable force field for use in molecular dynamics (MD) simulations. We investigate the use of multiple atoms-in-molecules (AIM) strategies to partition QM determined molecular electron densities into atomic subregions. The partitioned atomic densities are subsequently used to compute atomic dispersion coefficients from effective exchange-hole-dipole moment (XDM) calculations. In order to derive values for the repulsive van der Waals parameters from first principles, we use a simple volume relation to scale effective atomic radii. Explicit inclusion of higher order dispersion coefficients was tested for a series of alkanes, and we show that combining C_6 and C_8 attractive terms together with a C_{11} repulsive potential yields satisfying models when used in combination with our van der Waals parameters and electrostatic and bonded parameters as directly obtained from quantum calculations as well. This result highlights that explicit inclusion of higher order dispersion terms could be viable in simulation, and it suggests that currently available QM analysis methods allow for first-principles parametrization of molecular mechanics models.
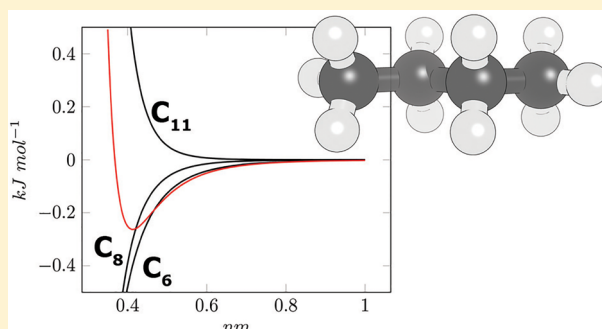
INTRODUCTION

Atomistic force fields are becoming increasingly important for explaining events in biological systems and for molecular design.[1−3] Through the use of GPU accelerated computing and specialized hardware it is now possible to simulate large biologically relevant systems on unprecedented time scales.[4,5] These simulations are reliant on the performance of the underlying force-field description of the relevant molecular interactions.[6,7] Next-generation atomistic force fields are polarizable force fields, which explicitly model electron density deformation by external electric fields stemming from environmental effects.[8,9] While still employing simple point potential functions, polarizable models can help improve our ability to describe typical properties of interest.[10]

Parametrizing molecules for force fields can be laborious, particularly for polarizable force fields, which contain additional parameters to describe the inducible dipole moments.[11] In addition most force-field optimizations are underdetermined by nature as they typically require calibrating too many parameters on a relatively small set of observables. The ability of biomolecular force fields to describe processes and equilibria of interest is usually measured by the ability to reproduce experimentally determined observables in condensed and hydrated phases.[7] Considering that the number of model compounds for which high-quality experimental data are available can be limited,[12] this optimization problem cannot be simply solved by a broad extension of calibration data sets.

As a consequence, minimizing the amount of free variables in the parameter optimization step can strongly accelerate force-field development. For that purpose bonded parameters describing bond stretching and angle bending have been directly and successfully derived in several automated parametrization protocols[12,13] from quantum mechanical (QM) Hessian calculations. Electrostatic parameters can be derived from a gas-phase ground-state calculation, either by fitting the molecular electrostatic potential on a grid around the compound resulting in electrostatic potential (ESP) derived charges[12−14] or by partitioning the computed electron density in an atoms-in-molecules (AIM) manner, where the difference between the effective number of electrons and the nuclear charge determines the partial charge of the atoms.[15−19] We recently showed that the additional electrostatic parameters in nonadditive force fields (polarizabilities of the inducible dipole moments) can be derived from combined QM/MM calculations in the condensed phase, and we successfully applied this approach in first parametrization studies.[20−22]

A complicating factor in many force-field calibration efforts is the optimization of van der Waals parameters, which are typically fitted based on condensed-phase properties of a training set of molecules. Applying response theory to determine values for the attractive C_6 dispersion parameter

for use in molecular simulation is computationally intensive and impractical for larger molecules.[23] Moreover a partitioning into atom−atom contributions from a molecular dispersion strength is required, which can be nontrivial. As a result most of the available force fields use tabulated van der Waals parameters that are assigned based on local atom typing.[7,24−27] While significant progress has been made in automated force-field parametrization through smart site typing[28−30] and automatic calibration through gradient descent,[20,31] covering a large part of possible chemical space for, e.g., drug-like compounds is still difficult for calibration-based approaches with many free parameters. However, recent work by Cole et al.[32,33] to determine dispersion parameters for classical force fields from atomic reference values scaled based on AIM derived atomic volumes shows promise as an ab initio approach to estimate attractive van der Waals parameters. Recently Mohebifar et al.[34] used the exchange-hole−dipole moment (XDM) model and density functional theory (DFT) calculations at the PBE0/aug-cc-pVTZ level of theory to determine dispersion constants of 88 compounds. These constants were found to be systematically and significantly lower than the values currently used in biomolecular force fields.[34−37] This raises the question of whether many of these force-field parameter sets are currently in a suboptimal part of parameter space simply because traditional calibration efforts work on relatively small databases of compounds (due to time and/or experimental data quality constraints), where the minimal number of free parameters in the calibration is in the order of the number of relevant independent experimental observations. To answer this question, several complications in parameter optimization have to be kept in mind. Higher order dispersion terms that comprise interactions involving higher order multipole moments can contribute up to more than 30% of the total dispersion energy and cannot be ignored.[35] However, it would be hard to justify adding additional dispersion terms and associated free parameters (C_8, C_{10}) when already facing an underdetermined calibration problem. As a result C_6 dispersion coefficients in force fields have to be large enough to compensate for the omission of higher order terms.[34,37] In addition, due to the omission of explicit polarization treatment, additive force fields partially encode polarization effects in their dispersion parameters. While the partial charge model in static-charge force fields is determined under polarizing conditions (e.g., from COSMO calculations[38]), a re-distribution of charges cannot fully envelop the full scope of polarization space (for example when using partial charges to describe a polarized flat molecule such as benzene). This leaves a potentially substantial role for explicitly polarizable models in further minimizing the number of free parameters in force-field calibration, as a more appropriate balance between electrostatic and dispersion forces could be achieved. A polarizable model can also aid in preventing possible inconsistencies when determining properties and parameters either in vacuum and/or condensed-phase environments. In the current work, nonbonded parameters are obtained from gas-phase electronic structure calculations, following the principle that any density shift under influence of electric fields should be (solely) modeled by explicitly treating electronic polarization. The set of (fixed) partial charges used in this study is hence determined in vacuo, preventing double counting of polarization contributions. When using a dielectric continuum with given dielectric permittivity (ϵ_{medium}) to model solvation, one should take into

account the self-polarization and distortion contributions, which cannot be represented well in a nonpolarizable force field. Therefore, as reported by Cole et al.,[32] the choice of ϵ_{medium} may not represent solvation properly but rather reflect a state in which the contribution of polarization and distortion cancel out, leading to effective inclusion of induction effects. Similarly, if a condensed-phase environment is described by including static Bq (MM) charges in the QM Hamiltonian, one should consider that the static charges from force fields are typically corrected for polarization and distortion effects and are therefore producing a too small dipole moment of the solvent molecules included in the quantum calculation.[39]

In this work we investigate the applicability of using a polarizable model that explicitly includes a higher order (C_8) dispersion term together with a force-field parameter set directly derived from quantum calculations, to achieve a physical description of dispersion effects while simultaneously minimizing the number of free variables in parameter optimization. In order to quantum mechanically determine van der Waals parameters, we partition computed molecular electron densities over the available atoms such that we can use the XDM model to estimate dispersion coefficients C_6 and C_8 from atomic-based densities (ρ_a):[36]

$$C_6 = \frac{\alpha_i \alpha_j \langle M_1^2 \rangle_i \langle M_1^2 \rangle_j}{\alpha_j \langle M_1^2 \rangle_i + \alpha_i \langle M_1^2 \rangle_j} \tag{1}$$

$$C_8 = \frac{3}{2} \frac{\alpha_i \alpha_j (\langle M_1^2 \rangle_i \langle M_2^2 \rangle_j + \langle M_2^2 \rangle_i \langle M_1^2 \rangle_j)}{\alpha_j \langle M_1^2 \rangle_i + \alpha_i \langle M_1^2 \rangle_j} \tag{2}$$

Calculating C_6 and C_8 for a pair of atoms i and j requires their multipole moment integral M_l^2 and their static atomic polarizability α_i or α_j (derived here as described below). Repulsive van der Waals parameters can be estimated quantum mechanically by using (tabulated) free and (computed) in-molecule atomic volumes obtained from QM calculations on isolated atoms and ρ_a, respectively, to determine values for AIM static polarizabilities α_i:[32,35]

$$\alpha_i = \frac{\langle r^3 \rangle_i}{\langle r^3 \rangle_{i,free}} \alpha_{i,free} \tag{3}$$

Note that this approximation to derive α_i is inherently built into XDM. Using the Slater−Kirkwood approximation that links polarizability to atomic radii (r_0),[40] Miller derived an approximation (eq 4) to compute effective van der Waals radii from polarizabilities,[41] where the prefactor is empirically derived to reproduce van der Waals radii tabulated by Bondi.[42]

$$r_{0,i} = 1.05 \sqrt{3} [\alpha_0 \alpha_i]^{1/4} \tag{4}$$

Atomic repulsive van der Waals parameters C_x for use in combination with C_6 and C_8 can be directly determined from the zero-energy point at the van der Waals potential energy function as employed by us, eq 5. Note that use of different values for the exponent x in the repulsive van der Waals term are evaluated in this work.

$$V_{vdW} = -\frac{C_6}{r^6} - \frac{C_8}{r^8} + \frac{C_x}{r^x} \tag{5}$$

For the atomic decomposition of QM determined electron densities, we rely here on AIM approaches to obtain ρ_a's from the molecular density. AIM methods assign each atom a weight

to each of the grid points in the underlying density integration grid, under the condition that the division of electron density on each grid point is complete.[18] All subsequent analyses can therefore be performed at an atomic level, as required for a fully atomistic force field. Popular choices for AIM partitioning schemes include the original Hirshfeld method,[15] in which weights are assigned to each atom (for each integration grid point) based on the local density stemming from a pro-atom placed at the nuclear site. A clear disadvantage of Hirshfeld paritioning is that initial pro-atomic charge states are arbitrarily set and may be wrongly assigned.[18] Iterative Hirshfeld (Hirshfeld-I)[16] introduces an iterative scheme to overcome this limitation and interpolates the value for the pro-atomic states by using a linear combination of the two pro-atoms with closest values for their net atomic charge (as compared to the value obtained in the last iteration). We also evaluated use of the iterative-stockholder analysis (ISA) scheme,[43] where instead of employing a predefined functional form of the pro-atom, a spherical average of the local atomic density is used. As an additional AIM method we tested use of mimimal-basis-iterative-stockholder (MBIS),[17] in which the pro-atom density is expressed as a series of s-type slater functions. A more detailed description of the AIM methods considered here is given in a recently published review by Heidar-Zadeh et al.[18] As part of our efforts we compare the use of these four AIM schemes to partition molecular electron densities into atomic ones. The performance of the different schemes in combination with XDM is evaluated for a series of alkane model compounds (listed in Table 1), which predominantly

Table 1. Simulation Conditions for the 11 Alkane Systems Considered: Values for the Reaction-Field Dielectric Permittivity (ϵ_0) and Simulation Temperature (T, in K; Based on the Measurement Temperature of the Density (ρ) and Heat of Vaporization (ΔH_{vap}))[a]

compound	ϵ_0	$T(H_{vap})$	$T(\rho)$
propane	1.80	231	231
butane	1.77	273	273
pentane	1.84	298	293
hexane	1.88	298	298
heptane	1.91	298	298
octane	1.95	298	298
nonane	1.96	298	298
isobutane	1.83	261	261
isopentane	1.85	298	293
3-methylpentane	1.89	298	298
3-ethylpentane	1.94	298	298

[a]Reference data and measurement temperature taken from ref 64 and 65.

interact via dispersion. Encouragingly, we find that an optimal part of model parameter space for force-field calibration can be directly derived from quantum calculations. Note that in order to determine charge populations and assign partial atomic charges for our final model, we found here a fifth AIM scheme to be most suitable, i.e., DDEC3.[44,45] In this scheme, core and valence electrons are handled separately, and therefore it provides the best performing set of atomic charges in this work. DDEC3 was only used for partial charge assignment here, as its introduction into a XDM dispersion analysis is currently not feasible yet.

METHODS

Polarizable System. An explicit polarizable model is used in the molecular dynamics (MD) simulations performed in this work. Inducible dipole moments were added to all atoms (including hydrogens) using the charge-on-spring (COS) model,[46,47] where positions of the massless virtual COS sites are iteratively optimized each time step in the simulation according to the local electric field, assuming isotropic and linear response. Virtual sites with a net charge of -8.0 e are used, neutralized by a counter charge of $+8.0$ e added to the polarizable atom in question. A Thole or other damping model is not used, and therefore all dipole interactions between 1,2 and 1,3 neighbors are excluded. During the local electric field calculations, 1,4 interactions are excluded as well, in accordance with the GROMOS implementation of the COS model.[46,48] Force constants of the springs connecting the COS and polarizable atoms are determined using atomic polar-izabilities documented by Miller, with values of sp[3] carbon and hydrogen atoms set to 1.061×10^{-3} and 0.387×10^{-3} nm^3, respectively.[41]

Type Annotations. To enable accounting for the effect of local chemical environments on values for force-field parameters while keeping our model simple and transferable, atoms with similar chemical environment were clustered into type annotations for all bonded, van der Waals, and static electrostatic parameters. For the atomic charges we considered a chemical environment up to two bonds away to be identical and encoded this into local environment subgraphs.[49] For dispersion parameters the local environment subgraphs were reduced to a single-bond neighbor graph, significantly reducing the number of van der Waals types considered. Angle and covalent bond type assignments were based on the specific elements under consideration and did not include additional environmental rules, but dihedral assignments used van der Waals types instead, resulting in an exhaustive set of dihedral parameters.

Bonded Parameters. For the determination of bonded parameters in an automated fashion, geometry optimizations were performed and Hessians were calculated at the B3LYP/aug-cc-pVTZ level of theory using NWChem 6.8,[50] for all alkane molecules under consideration. Harmonic force constants for bonds and angles were computed using a Hessian matrix projection according to the modified Seminario approach.[51] Zero-energy bond lengths and angles were derived directly from the geometry-optimized coordinate sets. Force constants and zero-energy bond length and angle values were averaged for each respective bonded type.

Because a completely new set of interaction parameters was used here, we needed to reoptimize the dihedral profiles. A GROMOS dihedral potential energy term was assigned to every dihedral type, with multiplicity and phase shift set to 3 and 0, respectively.[7] We performed dihedral potential energy surface scans at the B3LYP/aug-cc-pVTZ level of theory using GAMESS-US (version 2014) to determine the force constant associated with the amplitude.[52] A dispersion correction was introduced in these calculations using Grimme's D3 model.[53] For each dihedral scan, 15° angle increments were applied and optimized geometries from unconstrained geometry optimizations at the same level of theory were used as starting structure by prerotating to a given dihedral angle value. **Z**-matrix constraints were used to keep dihedral angles constant during dihedral-scan geometry optimizations. Force constants for each

dihedral were fitted by linear regression to the difference profile of the lowest and highest energy states, as obtained from the scan for the smallest alkane molecule in which the dihedral type was occurring.

XDM Calculations and Partial Charge Assignment. Gas-phase optimized coordinates (determined at the B3LYP/ 6-31G* level of theory) were downloaded from the ATB for a set of 11 small molecule training compounds.[12] Subsequently this starting geometry was refined using NWChem 6.8,[50] using a B3LYP functional and an augmented Dunning basis set aug-cc-pVTZ.[54,55] Because the postg program[56,57] was employed for the XDM postprocessing of the resulting densities, we stored them in Molden files and subsequently converted these to the postg recommended .wfx format using Molden2AIM.[56–58]

The published version of postg only provides the option of Hirshfeld partitioning for AIM analysis prior to the calculation of atomic exchange dipole holes.[15] Therefore, postg was modified to accept an external Becke grid with corresponding Hirshfeld weights. Horton version 2.1[59] was used for the density partitioning following the Hirshfeld, Hirshfeld-I,[16] ISA,[43] and MBIS methods.[17] Both Hirshfeld and Hirshfeld-I require a pro-atom database with electron densities of the isolated elements in multiple charge states. For consistency these were generated using the same functional (B3LYP) and aug-cc-pVTZ basis set as used in the molecular density calculations. Partial charges were directly derived from the difference between the effective number of electrons assigned during AIM partitioning and the nuclear charge, making the charge model consistent with the dispersion parameter assignment.

MD Simulations. Pure liquid MD simulations of the alkane compounds were performed using GROMOS11 (release 1.4.0),[48] which was modified to include higher order dispersion coefficients. Starting geometries for all molecules were taken from the ATB,[12] and 1024 molecules were packed into periodic cubic boxes using the GROMOS++ tool ran_box.[12,60] After steepest descent energy minimization, initial atomic velocities were generated based on a Maxwell–Boltzmann distribution at 50 K and the systems were subsequently slowly heated in five subsequent NVT simulations of 20 ps each to reach the target simulation temperature (Table 1). After this thermalization of the system, 1.5 ns of pre-equilibration time was used under production conditions. In all simulations, leap-frog integration was used to integrate the equations of motion with a 2 fs time step. Production simulations lasted in total 5 ns, and energies and coordinates were stored every 100 and 500 time steps, respectively. Constant temperature was maintained with a Berendsen weak coupling thermostat and a corresponding coupling constant of 0.1 ps.[61] For systems for which experimental densities (ρ) were not available at the same temperature as heats of varporization (ΔH_{vap}), independent simulations were run at each temperature. In the production (and preproduction) simulations, pressure was coupled weakly to 1 atm using a Berendsen barostat[61] with a 0.5 ps coupling constant and an isothermal compressibility set to 4.575×10^{-4} kJ mol^{-1} nm^{-3}. Center of mass motion of the systems was removed every 1000 time steps in a linear manner. All bond lengths were constrained using SHAKE with a relative geometric tolerance criterion of 10^{-4}.[62] Interaction energies and forces were computed using a charge group based cutoff and a multiple time step algorithm: they were evaluated every

time step for pairs of atoms within an inner cutoff of 0.8 nm, while contributions from pairs of atoms within an intermediate range (between 0.8 and 1.4 nm) were computed every five time steps. Long range electrostatic interactions beyond 1.4 nm were treated by the reaction-field approach, using dielectric constants for the pure liquids simulated as listed in Table 1.[63]

Simulations of single-model compounds in the gas phase were coupled to a stochastic bath to maintain the target simulation temperature. The internal friction coefficient was set to 24 ps^{-1}. Gas-phase production runs were in total 20 ns.

Equation 6 shows how for each alkane the enthalpy of vaporization (ΔH_{vap}) was computed using the averaged potential energy from a condensed phase (U_{con}^{pot}) and gas-phase simulation (U_{gas}^{pot}), where N_{mol}, R, and T are the number of molecules in simulation, gas constant, and temperature of the system, respectively. Average potential energies were extracted from the energy trajectories using GROMOS++ tool ene_ana.[60]

$$\Delta H_{vap} = U_{gas}^{pot}(T) - \frac{1}{N_{mol}} U_{con}^{pot}(T) + RT \tag{6}$$

■ RESULTS AND DISCUSSION

Choice of AIM Method. To quantum mechanically derive model parameters at the atomic level of detail, we used and compared four popular and/or state-of-the-art AIM schemes. The first method considered is Hirshfeld partioning,[15] in which pro-atom densities of the isolated atoms are used to (noniteratively) partition the molecular density into atomic contributions. It is known that the pro-atom choice in Hirshfeld partitioning typically leads to an underestimated charge separation.[18] Consequently the alkanes considered in this work show relatively small values for the individual atomic charges when employing the Hirshfeld partitioning scheme, with partial charges ranging between −0.087 and 0.027 e, Supporting Information Table S1. Similarly, the corresponding dispersion parameters calculated from XDM show a small spread for different carbon atom types (Table 2), consistent with results presented by Mohebifar et al.[34]

In the iterative-Hirshfeld scheme, the atomic-based electron densities are determined in an iterative approach in which atomic weights at the integration grid points are updated every iteration based on the results of the previous one.[18] This approach has been shown to result in an improved description of molecular electrostatic potentials.[18] With Hirshfeld-I we find distinctly different dispersion coefficients for the carbon atoms depending on the number of attached hydrogens, Table 2. A larger part of the hydrogen−carbon boundary region is assigned to the carbon atoms resulting in significantly larger estimates for the negative partial charges and the dispersion coefficients of the carbons with multiple hydrogen atoms attached, when compared to the values obtained using noniterative Hirshfeld (Tables 2 and S1).

Minimal-basis-iterative-stockholder was suggested to be an improvement over Hirshfeld in partitioning electron densities of oxides, where the dependency on a pro-atom database was removed.[17] While direct performance of XDM on MBIS partitioned densities has to our knowledge not been reported in literature, results from previous work in which the Tkatchenko−Scheffler method was used to derive dispersion coefficients[66] suggest that XDM may also benefit from using this relatively new partitioning method. With MBIS we find

Table 2. C_6 and C_8 Dispersion Constants Calculated Using Four Different AIM Methods (Hirshfeld, Hirshfeld-I, MBIS, and ISA) for the Four van der Waals Atom Types Considered in This Work[a]

	C_6	C_8	r_0
Hirshfeld			
HC	2.41	52.9	0.1275
CH1	18.49	675.1	0.1709
CH2	19.67	706.4	0.1698
CH3	21.11	748.0	0.1687
Hirshfeld-I			
HC	1.88	38.4	0.1234
CH1	16.88	594.4	0.1688
CH2	24.72	1050.3	0.1766
CH3	33.99	1574.2	0.1820
MBIS			
HC	0.99	14.3	0.1090
CH1	19.49	766.3	0.1714
CH2	26.82	1140.4	0.1766
CH3	37.65	1695.2	0.1820
ISA			
HC	2.17	44.3	0.1231
CH1	9.60	255.0	0.1516
CH2	13.30	363.4	0.1576
CH3	20.06	624.3	0.1654

[a]HC denotes a generic alkane hydrogen, while CHx denotes an sp^3 carbon with x hydrogens bound to it. Column r_0 lists predicted atomic radii derived from static polarizabilities using the Slater–Kirkwood approximation. C_6 and C_8 are given in atomic units; r_0 in nanometer.

slightly stronger charge separation between the carbon and hydrogen atoms when compared to Hirshfeld-I, resulting in larger partial charges for most atom types (Table S1) and smaller dispersion coefficients for the hydrogens (Table 2). We also tested the iterative-stockholder analysis partitioning scheme, which was developed independently from Hirshfeld-I. Table 2 shows that dispersion is highly underestimated by ISA compared to the other AIM partitioning schemes, probably due to the spherical averaged pro-atom densities used in ISA. Because of this discrepancy in the dispersion calculations, ISA was not considered further in deriving alkane models for the pure liquid simulations in the last part of this work.

Similarly to applying damping functions, the evaluation of van der Waals interactions relies on using atomic van der Waals radii, which describe the distance at which the van der Waals interaction energy between two particles is zero. Values for these radii are not trivial to derive.[35] Typically they have been calibrated against pure liquid reference data, but having well-defined starting points for calibration is preferential. For that purpose we used eqs 3 and 4 to quantum mechanically compute radii as starting points, in combination with the four different AIM partitioning schemes. The results are listed in Table 2 and show in general a smaller variation in the calculated radii among the different AIM schemes than observed for the computed dispersion parameters.

Pure Liquid Performance. To verify the use of an additional dispersion term and of the van der Waals parameters derived from the QM, AIM, and XDM calculations, we performed pure liquid simulations of alkanes because their properties predominantly depend on the strength of involved dispersion interactions. In MD simulations the repulsive part of

the van der Waals potential energy function is typically modeled with a C_{12} asymptotic function, chosen for computational efficiency. Here we opted to test which exponential term gave the best performance in combination with a C_6 and C_8 potential. We tested C_{11} to C_{14} repulsive functions (i.e., x in eq 5 = 11, 12, 13, or 14) for a first selection of four alkanes (propane, butane, hexane, and isobutane) and for each of the AIM methods under consideration, Tables S2 and S3. From all 12 models, Hirshfeld-I combined with a C_{11} repulsive scheme appears the most promising with a root-mean-square deviation (rmsd) of computed values from experiment of 21.4 kg m^3 and 1.15 kJ mol^{-1} for density (Table S2) and heat of vaporization (Table S3), respectively. While the C_{11} repulsive term is not commonly used in molecular simulation, there is a theoretical basis for this repulsive shape, as higher order dispersion calculations on hydrogen show that the odd indexed (e.g., C_{11} and C_{13}) are repulsive, stemming from third order perturbation theory.[67] As summarized in Table 3, Hirschfeld-I thus

Table 3. Pure-Liquid Simulation Results for a Set of Four Alkanes, Using Dispersion Parameters Obtained from XDM and Our Predicted van der Waals Radii[a]

	ρ(exp)	ρ(sim)	ΔH_{vap}(exp)	ΔH_{vap}(sim)
Hirshfeld (C_{12})				
propane	582.5	565.4	18.87	20.04
butane	600.7	564.6	22.4	22.89
hexane	654.8	606.2	31.6	30.11
isobutane	594.0	587.8	21.42	23.50
rmsd		31.6		1.43
Hirshfeld-I (C_{11})				
propane	582.5	576.4	18.87	20.08
butane	600.7	569.0	22.4	22.58
hexane	654.8	627.2	31.6	31.28
isobutane	594.0	599.3	21.42	23.35
rmsd		21.4		1.15
MBIS (C_{11})				
propane	582.5	646.7	18.87	20.55
butane	600.7	618.4	22.4	22.82
hexane	654.8	712.7	31.6	33.52
isobutane	594.0	684.2	21.42	24.59
rmsd		63.1		2.0

[a]For each of the AIM methods tested, use of the best performing repulsive exponential shape is listed (indicated in parentheses per AIM method). Experimental (exp)[64,65] and calculated values (sim) for densities (ρ) and heats of vaporization (ΔH_{vap}) are given in kg m^{-3} and kJ mol^{-1}, respectively, together with values for the root-mean-square deviation (rmsd) from experiment.

outperformed the other tested combinations (Table 3), in line with our finding that Hirschfeld-I performs well in predicting molecular C_6 values when compared to the corresponding experimental (DOSD oscillator) estimates,[68] Table S4.

Encouraged by the results for our set of four training compounds, we tested model performance for a larger range of alkanes, including multiple branched alkanes (Table S5). In general the computed values for ΔH_{vap} are close to experiment (with a maximum deviation of 1.9 kJ mol^{-1}), implying that we successfully predicted a region of parameter space with

balanced interaction energies. However, for several of the branched alkanes (isopentane, 3-methylpentane, and 3-ethylpentane), calculated densities are significantly off from experiment with a rmsd of 50 kg m^{-1}, Table S5. We traced this discrepancy to electrostatic interactions. The Hirshfeld-I model is dipole agnostic, and the resulting static molecular dipole moments were too large compared to their experimental estimates (approximately 0.3 D for the branched alkanes, while the value should be on the order of 0 D). Upon switching to the DDEC3 AIM method for determination of partial charges,[44,45] the in-molecule charge separations become less pronounced. In the DDEC3 scheme, core and valence electrons are explicit, which results in a more stable set of charges. Unsurprisingly our partial charges (Table S1) are similar to the ones used in the CHARMM Drude oscillator (DO) based polarizable force field for alkanes,[69] which were obtained from electrostatic potential based fitting and range from +0.095 to −0.177 e. For us, in terms of reproducing experimental values for density and heat of vaporization, DDEC3 was superior for charge determination over all other AIM methods tested. Using the DDEC3 derived point charge model, the results of our simulations are consistent with experiment (Table 4), with a rmsd from experiment of only 19.9 kg m^{-3} and 0.64 kJ mol^{-1} for the density and heat of vaporization, respectively.

Table 4. Pure Liquid Performance of Our Hirshfeld-I Decomposition Based Model Using a C_{11} Repulsive Shape and Partial Atomic Charges Obtained with the DDEC3 AIM Method[a]

compound	ρ(exp)	ρ(sim)	ΔH_{vap}(exp)	ΔH_{vap}(sim)
propane	582.5	574.8	18.87	19.60
butane	600.7	580.1	22.4	22.82
pentane	626.2	606.1	26.43	26.47
hexane	654.8	634.8	31.6	31.55
heptane	679.4	660.2	36.65	37.02
octane	698.6	679.7	41.49	42.01
nonane	713.8	695.0	46.4	46.45
iso-butane	594.0	599.2	21.42	23.1
isopentane	620.1	594.8	24.85	25.26
3-methylpentane	659.8	633.7	30.28	29.69
3-ethylpentane	693.8	668.2	35.22	35.30
rmsd		19.9		0.64

[a]Experimental (exp)[64,65] and calculated (sim) values for the densities and heats of vaporization are given in kg m^{-3} and kJ mol^{-1}, respectively.

Tables S6 and S7 compare the performance of our final model with other COS or Drude oscillator based polarizable force fields for hydrocarbons, i.e., the CHARMM-DO[69] and the GROMOS 45A3/COS alkane models.[70] Both were calibrated to reproduce experimental values for bulk liquid properties, and from a comparison with experiment, they perform slightly better than our model for the alkanes that were included both in this work and in ref 69 or 70 (Tables S6 and S7). We also find that both the CHARMM-DO and GROMOS 45A3/COS[70] force fields comprise effective C_6 parameters that are significantly larger than those calculated by XDM (cf. Table S8), as was reported earlier by Mohebifar et al.[34] In terms of fitted and calculated effective radii, our model utilizes the same part of parameter space as the CHARMM-DO force field, in particular for the CH$_2$ and CH$_3$ groups; see Table S8 as well.

■ CONCLUSIONS

In this work we present a strategy to develop an all-atom force field in which every parameter is determined from first principles, including higher order van der Waals constants. For a set of 11 linear and branched alkanes we found that by adding a single higher order dispersion term (C_8) to a polarizable model and using a C_{11} repulsive potential energy term, we reproduce experimental estimates for thermodynamic pure liquid properties of the alkanes with a rmsd of 19.9 kg m^3 and 0.64 kJ mol^{-1} for density and heat of vaporization, respectively, when using a set of model parameters that was completely derived from (a series of) quantum calculations. For that purpose we included application of the XDM model to compute dispersion coefficients and a Slater−Kirkwood approach to determine atomic radii. To derive AIM values for these radii and coefficients, we found the Hirschfeld-I partitioning scheme well-suited to obtain van der Waals parameters. In a next study we will examine whether XDM dispersion calculations have sufficient accuracy to correctly predict dispersion parameters for a large variety of molecules of interest, including other model compounds for typical biomolecular building blocks. The notion that our strategy can provide high-quality starting points required for automated calibration of novel force fields (e.g., by using automated calibration procedures[71]) is encouraging. Furthermore we show that it is possible to capture higher order dispersion in a relatively simple and efficient model, where required changes to force-field and simulation software are minimal. According to recent studies in literature, effects from higher order dispersion may be important for a correct model description of, for example, the folding of disordered protein domains, which is an additional motivation for including such a term in our force fields.[72,73] While our work clearly demonstrates that it is possible to capture higher order dispersion terms explicitly in molecular simulation, it is undeniable that classical force fields with only a C_6 dispersion term have been successful for years. We speculate that the reasons that C_6 attractive potentials can be successful are 4-fold: (i) there is no exact definition for atomic radii that should be used, and therefore effective radii can be slightly lowered to deepen well depths (ϵ_{ij}). (ii) C_6 values can be raised to effectively include higher order dispersion coefficients, as addressed by Mohebifar et al..[34] (iii) The choice of combination rules can have a large impact on effective well depths for heterogeneous atom pairs. In this work we used additive combination rules where the total van der Waals distance between two heterogeneous atoms is the sum of their invididual radii ($\sigma_{ij} = r_i + r_j$). However, in other force fields (e.g., GROMOS and OPLS) geometric combination rules for the repulsive part are used ($C_x = \sqrt{C_{i,x}C_{j,x}}$), which effectively lowers van der Waals distances (σ_{ij}) for heterogeneous atom pairs. Finally, (iv) the shape of the repulsive function is purely empirical, as exchange interactions cannot be mapped to simple classical mechanics functions. Having switched to C_{11} in this work provides additional damping of near- to medium-ranged interactions, thereby effectively reducing ϵ_{ij}. Similarly, choosing a C_{13} or higher repulsive potential would deepen effective well depths, compensating for the omission of higher order dispersion terms.

Here we showed that for dispersion driven systems inclusion of higher order dispersion terms can lead to efficient force-field parametrization. Therefore such an approach may be advantageous when both C_6 and C_8 can be calculated efficiently for the molecules of interest. In addition we speculate that by following a physical regime of parameter space for all atoms, we may create more coherent interactions between heterogeneous atom types, possibly reducing the number of special interaction parameters that are currently used in force fields.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.8b01105.

> Additional supporting tables, which contain results from supporting simulations, simulation parameters, and comparison with other force-field parameter sets (PDF)

■ AUTHOR INFORMATION

Corresponding Author
*(D.P.G.) E-mail: d.p.geerke@vu.nl.
ORCID ⓘ
Daan P. Geerke: 0000-0002-5262-6166
Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Weingarth, M.; Van Der Cruijsen, E. A.; Ostmeyer, J.; Lievestro, S.; Roux, B.; Baldus, M. Quantitative analysis of the water occupancy around the selectivity filter of a K+ channel in different gating modes. *J. Am. Chem. Soc.* **2014**, *136*, 2000−2007.

(2) Verhoef, D.; Visscher, K. M.; Vosmeer, C. R.; Cheung, K. L.; Reitsma, P. H.; Geerke, D. P.; Bos, M. H. Engineered factor Xa variants retain procoagulant activity independent of direct factor Xa inhibitors. *Nat. Commun.* **2017**, *8*, 528.

(3) Ganesan, A.; Coote, M. L.; Barakat, K. Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discovery Today* **2017**, *22*, 249−269.

(4) Shaw, D. E.; et al. Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*; IEEE Press: Piscataway, NJ, USA, 2014; pp 41−53, DOI: 10.1109/SC.2014.9.

(5) Stanley, N.; De Fabritiis, G. High throughput molecular dynamics for drug discovery. *In Silico Pharmacol.* **2015**, *3*, 3.

(6) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical free energy methods for drug discovery: progress and challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.

(7) Riniker, S. Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. *J. Chem. Inf. Model.* **2018**, *58*, 565−578.

(8) Albaugh, A.; et al. Advanced Potential Energy Surfaces for Molecular Simulation. *J. Phys. Chem. B* **2016**, *120*, 9811−9832.

(9) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116*, 4983−5013.

(10) Geerke, D. P.; Van Gunsteren, W. F. Calculation of the free energy of polarization: Quantifying the effect of explicitly treating electronic polarization on the transferability of force-field parameters. *J. Phys. Chem. B* **2007**, *111*, 6425−6436.

(11) Baker, C. M. Polarizable force fields for molecular dynamics simulations of biomolecules. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2015**, *5*, 241−254.

(12) Koziara, K. B.; Stroet, M.; Malde, A. K.; Mark, A. E. Testing and validation of the Automated Topology Builder (ATB) version 2.0: Prediction of hydration free enthalpies. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 221−233.

(13) Huang, J.; Lopes, P. E.; Roux, B.; MacKerell, A. D. Recent advances in polarizable force fields for macromolecules: Microsecond simulations of proteins using the classical drude oscillator model. *J. Phys. Chem. Lett.* **2014**, *5*, 3144−3150.

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

(15) Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chem. Acc.* **1977**, *44*, 129−138.

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

(17) Verstraelen, T.; Vandenbrande, S.; Heidar-Zadeh, F.; Vanduyfhuys, L.; Van Speybroeck, V.; Waroquier, M.; Ayers, P. W. Minimal Basis Iterative Stockholder: Atoms in Molecules for Force-Field Development. *J. Chem. Theory Comput.* **2016**, *12*, 3894−3912.

(18) Heidar-Zadeh, F.; Ayers, P. W.; Verstraelen, T.; Vinogradov, I.; Vöhringer-Martinez, E.; Bultinck, P. Information-Theoretic Approaches to Atoms-in-Molecules: Hirshfeld Family of Partitioning Schemes. *J. Phys. Chem. A* **2018**, *122*, 4219−4245.

(19) Bleiziffer, P.; Schaller, K.; Riniker, S. Machine Learning of Partial Charges Derived from High-Quality Quantum-Mechanical Calculations. *J. Chem. Inf. Model.* **2018**, *58*, 579−590.

(20) Visscher, K. M.; Vosmeer, C. R.; Luirink, R. A.; Geerke, D. P. A systematic approach to calibrate a transferable polarizable force field parameter set for primary alcohols. *J. Comput. Chem.* **2017**, *38*, 508−517.

(21) Vosmeer, C. R.; Kiewisch, K.; Keijzer, K.; Visscher, L.; Geerke, D. P. A comparison between QM/MM and QM/QM based fitting of condensed-phase atomic polarizabilities. *Phys. Chem. Chem. Phys.* **2014**, *16*, 17857−17862.

(22) Vosmeer, C. R.; Rustenburg, A. S.; Rice, J. E.; Horn, H. W.; Swope, W. C.; Geerke, D. P. QM/MM-based fitting of atomic polarizabilities for use in condensed-phase biomolecular simulation. *J. Chem. Theory Comput.* **2012**, *8*, 3839−3853.

(23) Fransson, T.; Rehn, D. R.; Dreuw, A.; Norman, P. Static polarizabilities and C6 dispersion coefficients using the algebraic-diagrammatic construction scheme for the complex polarization propagator. *J. Chem. Phys.* **2017**, *146*, 094301.

(24) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.

(25) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 3499−3509.

(26) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.

(27) Boulanger, E.; Huang, L.; Rupakheti, C.; MacKerell, A. D.; Roux, B. Optimized Lennard-Jones Parameters for Druglike Small Molecules. *J. Chem. Theory Comput.* **2018**, *14*, 3121−3131.

(28) Huang, L.; Roux, B. Automated Force Field Parameterization for Nonpolarizable and Polarizable Atomic Models Based on Ab Initio Target Data. *J. Chem. Theory Comput.* **2013**, *9*, 3543−3556.

(29) Mobley, D.; Bannan, C. C.; Rizzi, A.; Bayly, C. I.; Chodera, J. D.; Lim, V. T.; Lim, N. M.; Beauchamp, K. A.; Shirts, M. R.; Gilson, M. K.; Eastman, P. K. Open Force Field Consortium: Escaping atom types using direct chemical perception with SMIRNOFF v0.1. 2018, *bioRxiv*, DOI: 10.1101/286542.

(30) Zanette, C.; Bannan, C. C.; Bayly, C. I.; Fass, J.; Gilson, M.; Shirts, M. R.; Chodera, J.; Mobley, D. Toward Learned Chemical Perception of Force Field Typing Rules. *J. Chem. Theory Comput.* **2019**, *15*, 402−423.

(31) 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.

(32) 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.

(33) Horton, J.; Allen, A.; Dodda, L.; Cole, D. QUBEKit: Automating the Derivation of Force Field Parameters from Quantum Mechanics. 2018. *ChemRxiv*, https://chemrxiv.org/articles/QUBEKit_Automating_the_Derivation_of_Force_Field_Parameters_from_Quantum_Mechanics/7247045.

(34) Mohebifar, M.; Johnson, E. R.; Rowley, C. N. Evaluating Force-Field London Dispersion Coefficients Using the Exchange-Hole Dipole Moment Model. *J. Chem. Theory Comput.* **2017**, *13*, 6146−6157.

(35) 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.

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

(37) Walters, E. T.; Mohebifar, M.; Johnson, E. R.; Rowley, C. N. Evaluating the London Dispersion Coefficients of Protein Force Fields Using the Exchange-Hole Dipole Moment Model. *J. Phys. Chem. B* **2018**, *122*, 6690−6701.

(38) Klamt, A.; Schuurmann, G. Cosmo - a New Approach To Dielectric Screening in Solvents With Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799−805.

(39) Cerutti, D. S.; Rice, J. E.; Swope, W. C.; Case, D. A. Derivation of Fixed Partial Charges for Amino Acids Accommodating a Specific Water Model and Implicit Polarization. *J. Phys. Chem. B* **2013**, *117*, 2328−2338.

(40) Slater, J.; Kirkwood, J. The Van Der Waals Forces in Gases. *Phys. Rev.* **1931**, *37*, 682−697.

(41) Miller, K. J. Additivity methods in molecular polarizability Additivity Methods in Molecular Polarizability. *J. Am. Chem. Soc.* **1990**, *112*, 8533−8542.

(42) Bondi, A. Van der waals volumes and radii. *J. Phys. Chem.* **1964**, *68*, 441−451.

(43) Lillestolen, T. C.; Wheatley, R. J. Atomic charge densities generated using an iterative stockholder procedure. *J. Chem. Phys.* **2009**, *131*, 144101.

(44) Manz, T. A.; Sholl, D. S. Chemically Meaningful Atomic Charges That Reproduce the Electrostatic Potential in Periodic and Nonperiodic Materials. *J. Chem. Theory Comput.* **2010**, *6*, 2455−2468.

(45) Manz, T. A.; Sholl, D. S. Improved atoms-in-molecule charge partitioning functional for simultaneously reproducing the electrostatic potential and chemical states in periodic and nonperiodic materials. *J. Chem. Theory Comput.* **2012**, *8*, 2844−2867.

(46) Yu, H.; Hansson, T.; Van Gunsteren, W. F. Development of a simple, self-consistent polarizable model for liquid water. *J. Chem. Phys.* **2003**, *118*, 221−234.

(47) Drude, P. *The theory of optics*; Longmans, Green: New York, 1902; p 592.

(48) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; Van Gunsteren, W. F. Architecture, implementation and parallelisa-tion of the GROMOS software for biomolecular simulation. *Comput. Phys. Commun.* **2012**, *183*, 890−903.

(49) Veen, L.; Geerke, D. P.; Mark, A. E.; Klau, G. W. Multiple-Choice Knapsack for Assigning Partial Atomic Charges in Drug-Like Molecules. *18th International Workshop on Algorithms in Bioinformatics (WABI 2018)*, Aug. 20−24, 2018, Helsinki, Finland; Schloss Dagstuhl: Dagstuhl, Germany, 2018; Vol. *113*, pp 1−13.

(50) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; De Jong, W. A. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **2010**, *181*, 1477−1489.

(51) Allen, A. E.; Payne, M. C.; Cole, D. J. Harmonic Force Constants for Molecular Mechanics Force Fields via Hessian Matrix Projection. *J. Chem. Theory Comput.* **2018**, *14*, 274−281.

(52) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; SU, S.; Windus, T.; Dupuis, M.; Montgomery, J. A. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993**, *14*, 1347−1363.

(53) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(54) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(55) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.

(56) Kannemann, F. O.; Becke, a. D. van der Waals Interactions in Density-Functional Theory: Intermolecular Complexes. *J. Chem. Theory Comput.* **2010**, *6*, 1081−1088.

(57) Otero-De-La-Roza, A.; Johnson, E. R. Non-covalent inter-actions and thermochemistry using XDM-corrected hybrid and range-separated hybrid density functionals. *J. Chem. Phys.* **2013**, *138*, 204109.

(58) Zou, W.; Nori-Shargh, D.; Boggs, J. E. On the covalent character of rare gas bonding interactions: A new kind of weak interaction. *J. Phys. Chem. A* **2013**, *117*, 207−212.

(59) Verstraelen, T.; Tecmer, P.; Heidar-Zadeh, F.; González-Espinoza, C. E.; Chan, M.; Kim, T. D.; Boguslawski, K.; Fias, S.; Vandenbrande, S.; Berrocal, D.; Ayers, P. W. *HORTON 2.1.0*; 2017; http://theochem.github.com/horton/.

(60) Eichenberger, A. P.; Allison, J. R.; Dolenc, J.; Geerke, D. P.; Horta, B. A.; Meier, K.; Oostenbrink, C.; Schmid, N.; Steiner, D.; Wang, D.; Van Gunsteren, W. F. GROMOS++ software for the analysis of biomolecular simulation trajectories. *J. Chem. Theory Comput.* **2011**, *7*, 3379−3390.

(61) Berendsen, H. J.; Postma, J. P.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(62) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(63) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451.

(64) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comput. Chem.* **2001**, *22*, 1205−1218.

(65) Lide, D. R., Ed. *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, FL, USA, 2005.

(66) Tkatchenko, A.; Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, *102*, 073005.

(67) Mitroy, J.; Bromley, M. W. Higher-order Cn dispersion coefficients for hydrogen. *Phys. Rev. A: At., Mol., Opt. Phys.* **2005**, *71*, 032709.

(68) Jiemchooroj, A.; Sernelius, B. E.; Norman, P. C6 dipole-dipole dispersion coefficients for the n-alkanes: Test of an additivity procedure. *Phys. Rev. A: At., Mol., Opt. Phys.* **2004**, *69*, 044701.

(69) Vorobyov, I. V.; Anisimov, V. M.; MacKerell, A. D. Polarizable empirical force field for alkanes based on the classical Drude oscillator model. *J. Phys. Chem. B* **2005**, *109*, 18988−18999.

(70) Szklarczyk, O. M.; Bachmann, S. J.; Van Gunsteren, W. F. A polarizable empirical force field for molecular dynamics simulation of liquid hydrocarbons. *J. Comput. Chem.* **2014**, *35*, 789−801.

(71) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable atomic multipole-based AMOEBA force field for proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046−4063.

(72) Robustelli, P.; Piana, S.; Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E4758−E4766.

(73) 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.