JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Density-Functional Theory with Dispersion-Correcting Potentials for Methane: Bridging the Efficiency and Accuracy Gap between High-Level Wave Function and Classical Molecular Mechanics Methods

Edmanuel Torres[†,‡] and Gino A. DiLabio*[,†,§]

[†]National Institute for Nanotechnology, National Research Council of Canada, 11421 Saskatchewan Drive, Edmonton, Alberta, Canada T6G 2M9

[‡]Faculty of Basic Sciences, Universidad Technológica de Bolívar, Cartagena, Colombia

[§]Department of Physics, University of Alberta, Edmonton Alberta, Canada T6G 2E1

Supporting Information

ABSTRACT: Large clusters of noncovalently bonded molecules can only be efficiently modeled by classical mechanics simulations. One prominent challenge associated with this approach is obtaining force-field parameters that accurately describe noncovalent interactions. High-level correlated wave function methods, such as CCSD(T), are capable of correctly predicting noncovalent interactions, and are widely used to produce reference data. However, high-level correlated methods are generally too computationally costly to generate the critical reference data required for good force-field parameter development. In this work we present an approach to generate Lennard-Jones force-field parameters to accurately account for noncovalent interactions. We propose the use of a computational step that is intermediate to CCSD(T) and classical molecular mechanics, that can bridge the accuracy and computational efficiency gap between them, and demonstrate the efficacy of our approach with methane clusters. On the basis of CCSD(T)-level binding energy data for a small set of methane clusters, we develop methane-specific, atom-centered, dispersion-correcting potentials (DCPs) for use with the PBE0 density-functional and 6-31+G(d,p) basis sets. We then use the PBE0-DCP approach to compute a detailed map of the interaction forces associated with the removal of a single methane molecule from a cluster of eight methane molecules and use this map to optimize the Lennard-Jones parameters for methane. The quality of the binding energies obtained by the Lennard-Jones parameters we obtained is assessed on a set of methane clusters containing from 2 to 40 molecules. Our Lennard-Jones parameters, used in combination with the intramolecular parameters of the CHARMM force field, are found to closely reproduce the results of our dispersion-corrected density-functional calculations. The approach outlined can be used to develop Lennard-Jones parameters for any kind of molecular system.

INTRODUCTION

The ability to model structures and properties of molecular clusters is critical to advancing important technologies related to, for example, gas capture (e.g., CO_2 sequesterization, H_2 storage),[1] gas-hydrates,[2] and liquid crystals.[3] The central challenge associated with modeling these, and other molecular aggregates, is that the structures are formed through van der Waals interactions, which are notoriously difficult to accurately reproduce using most simulation techniques. One exception to this is complete basis set extrapolated coupled cluster with single and double excitations, with the noniterative inclusion of corrections for triple excitations, i.e. CCSD(T)/CBS. This high-level correlated quantum mechanical wave function technique has been proven to be capable of predicting noncovalent interactions with great accuracy. However, this method has limited applicability because it requires significant computational time and resources to treat even the smallest systems.

Empirical force-field (FF) methods are routinely applied to very large chemical systems. For systems interacting through noncovalent forces, for example drug−protein active site interactions, FF methods are able to generate useful information about binding strengths related to enzyme inhibition, etc. FFs generally represent most van der Waals interactions through the use of a Lennard-Jones (L-J) or Morse like potential.[4] For example, the OPLS force-field uses an L-J expression in which the well-depths and collision diameters for interacting molecules are taken from the geometric mean of atomic parameters for individual atoms.[5] The parameters for the OPLS FF were optimized such that Monte Carlo simulations of 36 organic liquids reproduced experimental densities and heats of vaporization. Additional consideration was given to reproducing the complexation energies of neutral

Received: December 11, 2012

organic species with water as determined by low-level quantum theory (viz., HF/6-31G(d)). While FFs such as OPLS perform extremely well on the systems for which they were designed, and have been proven to be robust,[6] they nevertheless tend to perform poorly when applied to molecular clusters or crystals. Of course, FFs can be reparameterized in order to better treat specific, noncovalently interacting systems, but this often requires the use of time-consuming highly correlated wave function techniques, such as that described above.

Computational methods based on density-functional theory (DFT) have the potential to provide a bridge between highly correlated wave function methods and FF techniques. A critical deficiency of conventional DFT techniques is their well-known inability to accurately and reliably model noncovalent (especially dispersion) interactions. Several general approaches to correcting DFT methods for dispersion have been developed. For example, early efforts focused on the use of empirical, pairwise dispersion parameters in conjunction with density functionals,[7,8] an approach that was later popularized and expanded through the efforts of Grimme's group.[9−11] An elegant and particularly accurate approach to this was developed by Becke and Johnson, whose method calculates dispersion coefficients from the dipole moment of the exchange-hole.[12−14] Recently, Tkachenko and Scheffler developed a parameter-free approach to obtaining van der Waals parameters[15] that can be used to provide very good noncovalent binding energies.[16] Our own solution to the DFT "dispersion problem" is atom-centered dispersion-correcting potentials (DCPs), which improve the long-range behavior of the functional for which they were developed.[17] The DCP approach is similar in nature to a plane wave technique introduced by von Lilienfeld and co-workers.[18−20] Although all of these approaches were developed to generally improve the performance of DFT functionals for noncovalently bonded systems, their performances for specific systems are not, in all cases, accurate enough to provide reference data for FF development.[21]

High accuracy in dispersion-corrected DFT-based computational methods for application to noncovalent interactions can be achieved by developing corrections for specific systems or families of systems. DCPs are amenable to developing and applying accurate dispersion corrected DFT for specific systems for three key reasons: (1) DCPs are implemented by modifying the input file for standard computational packages, rather than by reprogramming, making their development and use straightforward. (2) Each of the Gaussian functions that constitute a particular DCP can be optimized to reproduce interactions that are particular to specific systems. This provides a great deal of flexibility in terms of describing noncovalent interactions. (3) DCPs can simultaneously accurately reproduce noncovalent interactions over large ranges of a potential energy surface (e.g., from minimum to dimer dissociation) while compensating for errors due to limitations in the size of the basis sets employed. This latter property of DCPs makes them particularly attractive for system-specific applications because this opens the possibility to the treatment of very large systems using small basis sets, allowing for the generation of detailed, accurate force maps and potential energy surfaces (PES), and/or realizing enormous savings in computational resources.

In this paper, we describe a system-specific approach that combines the utility and ease of use of DCPs with the broad applicability of DFT to bridge the gap between high-level correlated wave function theory and FF development. Using dispersion-bound methane clusters as case study species,[22] we develop hydrogen and carbon DCPs for use with PBE0[23]/6-31+G(d,p), without corrections for basis set incompleteness, that reproduce with high-fidelity one-dimensional, CCSD(T)/CBS PES for several methane dimers and the binding energies for several methane clusters. We name these system-specific dispersion-correcting potentials DCP(methane). The approach presented herein is general and adaptable to any DFT functional/basis set combinations. We then used the PBE0-DCP(methane)/6-31+G(d,p) approach to generate the forces associated with the interaction of a methane within a small methane cluster. These forces were subsequently used to derive new carbon and hydrogen L-J parameters, which were thereafter used to simulate methane clusters containing up to 40 molecules. The data we obtained are discussed in the context of a conventional and a specialized FF approach, and prospects for future applications are offered.

THEORETICAL BACKGROUND AND COMPUTATIONAL APPROACHES

DCP Development. DCPs are atom-centered potentials ($U(r)$) composed of Gaussian-type functions, as described in ref 24, with the following general form:

$$U_l(r) = r^{-2} \sum_{i=1}^{N_l} c_{li} r^{n_{li}} e^{-\zeta_{li} r^2} \quad (1)$$

In eq 1, l is the angular moment, N_l is the number of Gaussian functions, n_{li} is the power of r (set to 2 throughout this work), c_{li} is the coefficient of the Gaussian, and ξ_{li} is its exponent.

Earlier generations of DCPs were developed for the carbon and silicon atoms,[25] and corrected long-range interactions through the use of high-angular-momentum functions. More recently, we found that the use of lower angular momentum functions offered greater flexibility and the opportunity for unprecedented accuracy in the modeling of noncovalently bonded systems.[17] These functions act most strongly on the electron density in specific angular momentum channels. In this work, we develop alkane-specific DCPs composed of functions with s and p angular momentum for the H atom and s, p, d, and f angular momentum for the C atom.

The DCPs presented in this work were developed by optimizing c_{li} values (eq 1) associated with ξ_{li} values selected to span a range of values from 0.006 to 0.12. In certain cases, ξ_{li} values were also optimized. The optimizations were performed through the minimization of the error in the predicted interactions along one-dimensional PES shown in Figure 1. The PES were calculated by varying the distance between rigid, B3LYP/6-31G(d)-optimized monomers and computing the reference binding energies by performing CCSD(T)/CBS calculations according to an approach we used elsewhere.[26,27] The PES for three methane dimer conformations, having D_{3d}, C_{3v}, and D_{3h} symmetry, comprised our DCP "fitting" data set. Also included in the "fitting" set were the binding energies for a methane trimer and a methane tetramer.[28] These species were included to ensure that the optimized DCPs are capable of correctly reproducing many-body interactions associated with the noncovalently bonded methane clusters.[29,30] The full data associated with the DCP fitting are provided in the Supporting Information (SI). The Gaussian-03[31] and Gaussian-09[32] packages were used for the CCSD(T)/CBS calculations.

A set of DCPs were developed for hydrogen and carbon with the goal of reproducing the high-level, CCSD(T)/CBS data

dx.doi.org/10.1021/ct4003114 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

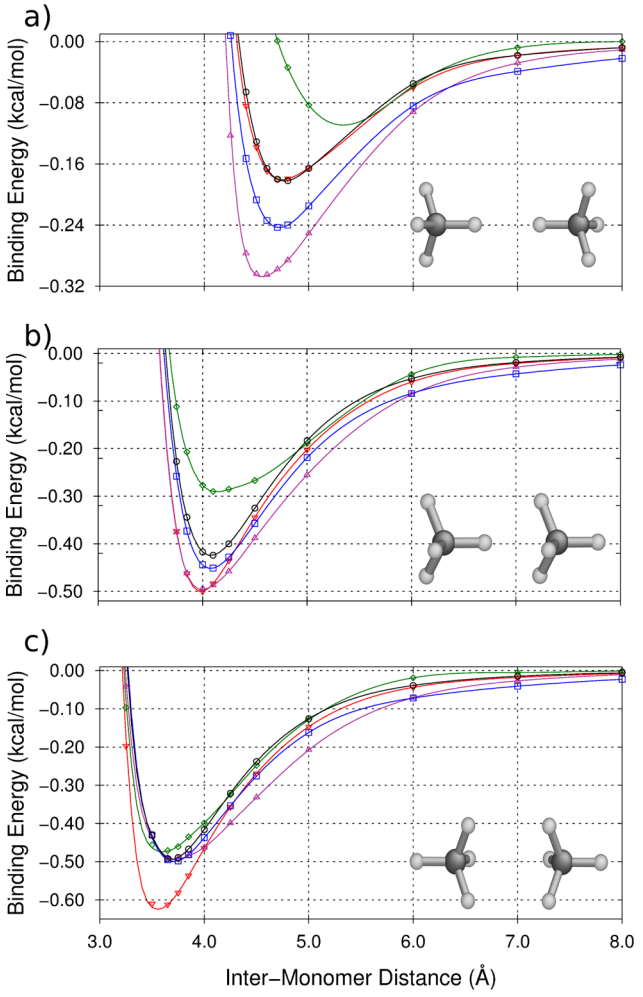

Figure 1. Potential energy surfaces for methane dimers using PBE0-DCP(methane)/6-31+G(d,p) (open squares). (a) D_{3h}; (b) C_{3v}; and (c) D_{3d} methane dimers. For comparison, PES obtained without counterpoise-corrected ω-B97X-D (down triangle), B97D (up triangles), M06-2X (diamonds), and CCSD(T)/CBS (circles) PES are also shown.

shown in Figure 1 and the BEs for methane trimer and tetramer. We chose to use the PBE0[33] functional in this work primarily because our initial efforts showed that well-performing DCPs were easily obtained with this functional. However, other functionals may be used in a similar fashion. For all of the PBE0-DCP calculations, we utilized 6-31+G(d,p) basis sets. These were selected because they possess enough flexibility, along with the DCPs, to accurately represent the predominately dispersion binding that occurs in methane dimers. These small basis sets offer a great deal of computational efficiency but are somewhat[34] incomplete. Therefore, the optimized DCPs must also compensate to some extent for errors arising from limitations in the basis sets employed while reproducing dispersion.

In order to generate the DCPs, we performed a stochastic search in the parameter space defined by the c_{li} and ξ_{li} values associated with the DCP functions (four for H and seven for C, vide infra) so as to minimize the mean absolute error (MAE) of the calculated binding energies in the fitting set according to eq 2:

$$\mathrm{MAE} = \frac{1}{M}\sum_{1}^{M}\left(\frac{1}{N}\sum_{1}^{N}|BE_i^{HL} - BE_i^{DCP}|\right) \tag{2}$$

Here, N accounts for the number of points associated with each of the PES in Figure 1, M is the number of PES, plus the single-points for the methane trimer and tetramer, and BE^{HL} and BE^{DCP} represent the binding energies calculated using the high-level approach and DCPs, respectively. To avoid getting trapped in local minima associated with the DCP parameters, the error was allowed to increase a small amount for up to three consecutive steps; otherwise, the optimization was restarted with slightly modified DCP parameters.

The DCP optimizations were performed using in-house scripts, which handled job submissions and data processing. Gaussian-09 was used to compute DFT energies in all cases. Because DCPs resemble effective core potentials, they can be used directly by the Gaussian package (and other computational chemistry programs) through a simple modification of input files. A utility program for generating Gaussian and NWChem input files with DCPs is available on our Web site,[35] and some input examples are provided in the Supporting Information.

For comparison, we also calculated the methane dimer PES using a variety of commonly available dispersion-corrected DFT methods. These results were also obtained using the Gaussian-09 program.

Force-Field Parameter Development. We considered a number of different approaches for using PBE0-DCP-(methane)/6-31+G(d,p) to generate L-J parameters for methane. We ultimately settled upon a two-step approach. An initial optimization of parameters was performed with the goal of minimizing the errors in predicted BEs along the methane dimer PES. The parameters obtained from this effort were then used as the starting point for fitting the forces calculated for each of 35 structures obtained from the stepwise extraction of a single methane molecule from the minimum energy structure of a cluster of eight methane molecules. The geometry of the methane 8-mer was taken from the work of Takeuchi[36] and reoptimized using PBE0-DCP(methane)/6-31+G(d,p). The L-J potential is defined by

$$U(r_{ij}) = 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{3}$$

where ε is the well-depth and σ is the collision diameter.

In both steps of the optimization, ε and σ for H and C atoms were optimized. All the molecular mechanics calculations were performed with the DL_POLY classic package.[37] Throughout this portion of our work, we used point charges and intramolecular force constants from the CHARMM force field. We tested three mixing rules[38] for the cross interactions between H and C atoms in the generation of the L-J parameters, defined as follows:

Lorentz−Berthelot:

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ij} + \sigma_{ji})$$

$$\varepsilon_{ij} = (\varepsilon_{ii}\varepsilon_{jj})^{1/2} \tag{4}$$

dx.doi.org/10.1021/ct4003114 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Halgren HHG:

$$\sigma_{ij} = \frac{\sigma_{ii}^3 + \sigma_{jj}^3}{\sigma_{ii}^3 + \sigma_{jj}^2}$$

$$\varepsilon_{ij} = \frac{4\varepsilon_{ii}\varepsilon_{jj}}{(\varepsilon_{ii}^{1/2} + \varepsilon_{jj}^{1/2})^2} \quad (5)$$

Waldman−Hagler:

$$\sigma_{ij} = \left(\frac{\sigma_{ii}^6 + \sigma_{jj}^6}{2}\right)^{1/6}$$

$$\varepsilon_{ij} = (\varepsilon_{ii}\varepsilon_{jj})^{1/2}\left(\frac{\sigma_{ii}^3\sigma_{jj}^3}{\sigma_{ij}^6}\right) \quad (6)$$

RESULTS AND DISCUSSION

The DCPs optimized for use with PBE0/6-31+G(d,p) on methane clusters are listed in Table 1. The optimization procedure was relatively straightforward and presented no particular difficulties.

Table 1. Methane-Specific Hydrogen and Carbon Dispersion-Correcting Potentials (DCPs) Optimized for Use with the PBE0/6-31+G(d,p) Density-Functional Theory-Based Method[a]

atom	function type	ζ_i	c_i
H	P and higher	0.120 000 000	0.000 236 004
		0.090 000 000	−0.000 037 158
		0.006 000 000	−0.000 000 109
	S−P	0.120 000 000	−0.000 038 997
C	F and higher	0.139 810 621	0.000 005 319
		0.090 000 000	0.000 178 855
		0.040 122 758	−0.000 001 236
		0.010 000 000	−0.000 016 268
	S−F	0.090 000 000	−0.000 974 288
	P−F	0.060 000 000	−0.000 087 240
	D−F	0.050 000 000	−0.003 997 556

[a]Additional formatting is required in order to use these DCPs in computational chemistry programs; see the Supporting Information.

It should be stressed that the DCP(methane) functions are highly specific to the chemical system, i.e. methane, and the method, i.e. PBE0/6-31+G(d,p), for which they were developed. Their application to other hydrocarbon species (for example, longer-chain alkanes, alkenes, aromatics) or with counterpoise corrections[39] or other basis sets will produce results of poor quality.

The performance of the methane-specific DCPs, compared to CCSD(T)/CBS, is illustrated for the three methane dimer cases in Figure 1 and in more detail in the Supporting Information. Figure 1 compares the performance of the PBE0-DCP(methane)/6-31+G(d,p), along with other DFT approximations designed to treat noncovalent interactions (without counterpoise corrections).

From Figure 1 we see that all of the dispersion-corrected DFT approaches show rather variable performance. For example, ωB97XD[40] almost perfectly reproduces the PES surface associated with the methane dimer shown in Figure 1a but overbinds those in Figure 1b and c. M06-2X,[41] however, underbinds in Figure 1a and b but almost perfectly reproduces the high-level PES in Figure 1c. B97D[9] displays overbinding in Figure 1a and b and poor dissociation behavior in Figure 1c.

Our PBE0-DCP(methane) approach displays improved behavior in going from Figure 1a to c. For example, at its minimum the PES in Figure 1a is overbound by ca. 0.05 kcal/mol, and this overbinding is reduced to less than ca. 0.02 and 0.01 kcal/mol for analogous points on the PES in Figure 1b and c, respectively. This order of performance mirrors the increasing importance of the PES in terms of their contribution to the binding that occurs in clusters. That is, the types of atom−atom contacts that are represented at the minimum in the PES in Figure 1a are not expected to be relevant in methane clusters, but those at the minimum of the PES shown in Figure 1c are expected to play a more important role in methane clusters. It should be pointed out that PBE0/6-31+G(d,p) without the benefit of DCPs predicts all three curves in Figure 1 to be only weakly bound—see the Supporting Information.

In the earliest generations of DCPs,[25,42,43] proper accounting of long-range interactions between noncovalently bonded systems (i.e., dissociation behavior) was not made and DCPs tended to act near the minima of the PES describing the interactions. However, later generations of DCPs,[17,44] including those presented in this work, optimize Gaussian functions that span the range of space from distances below PES minima out to complete monomer separation. The PBE0-DCP(methane) energy curves illustrated in Figure 1 demonstrate how well DCPs reproduce in particular the correct dissociation behavior of methane dimers.

Table 2. Reference Binding Energies for Methane Clusters Containing from 2 to 6 Monomers and Signed Deviations of Binding Energies Predicted by Various Dispersion-Corrected Density-Functional Theory Methods[a]

	Me_n				
method	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
CCSD(T)/CBS[b]	0.525	1.344	2.393	3.560	4.924
PBE0-DCP(methane)[c]	−0.002	0.000	0.002	0.007	−0.005
ωB97XD[c]	−0.107	−0.333	−0.544	−0.971	−1.331
M06-2X[c]	−0.025	0.003	0.307	0.089	0.217
B97D[c]	0.038	0.072	−0.042	0.081	0.086
B3LYP-DCP[e]	0.153	0.407	0.686	1.012	1.397
B3LYP-XDM[d]	0.045	0.104	0.163	0.160	0.214
B3LYP-D3[d]	0.155	0.239	0.164	0.185	0.111

[a]All values are in kcal/mol. [b]See ref 28. [c]6-31+G(d,p) basis sets. [d]aug-cc-pVTZ. [e]6-31+(G2d,2p). Relative energies are DFT − CCSD(T).

dx.doi.org/10.1021/ct4003114 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

In Table 2, we list the CCSD(T)/CBS BEs predicted for clusters containing from 2 to 6 methane molecules, along with the signed errors in the BEs predicted by various DFT-based methods. For all methods, the same structures were used for the BE calculations, and these are provided in the Supporting Information. We begin our discussion of the data in Table 2 by examining the performance of some common dispersion-corrected DFT-based methods for their ability to predict the BEs in the small methane clusters. The ωB97XD functional performance becomes increasingly poor as the size of the clusters increases. Interestingly, the errors in the predicted BEs for clusters with $n > 2$ do not correspond to the number of pairwise interactions times the error in BE for $n = 2$ (viz., 1, 3, 6, 10, and 20 for $n = 2-6$, respectively). This suggests that the many-body interactions predicted by ωB97XD, whether they are correct or not, are non-negligible.

The M06-2X functional appears to be somewhat better behaved than the ωB97XD, giving small errors in the BEs for $n = 2$, 3, 5 and slightly larger errors for $n = 4$, 6. The absence of a clear trend in predicted BEs with cluster size prevents M06-2X from being applied with confidence to larger clusters. On the other hand, the good performance of B97D on the clusters listed in Table 2 suggests that this approach may be suitable for treating larger clusters. The use of the small basis set in this case (as opposed to the typically large basis sets employed for the optimization of the B97D approach) contributes to some extent to the good performance of the method.

Table 2 also contains BE data obtained using the B3LYP functional in combination with three different dispersion-correction approaches. The first is the B3LYP-DCP approach, which was developed by us to be a generally applicable method for noncovalently bonded systems.[17] Displaying similarly poor performance to ωB97XD, the B3LYP-DCP approach gives an inaccurate BE for the methane dimer, and this results in a very low performance on the larger methane clusters. Therefore, B3LYP-DCP is essentially unusable for methane clusters. On the other hand, both the XDM[12,13] and D3[11,45]-corrected DFT approaches perform fairly well for the clusters and may find general applicability to the modeling of methane clusters.

The performance of the PBE0-DCP(methane) approach in terms of its ability to predict the BE values of a set of small methane clusters (Me_n, $n = 2-6$) is exceptional. It is important to reiterate that the clusters with $n = 3$ and 4 were explicitly included in the DCP fitting, while that for $n = 2$ was implied to some extent by the D_{3d} dimer shown in Figure 1c. The data in the table show that the methane-specific DCPs reproduces the high-level BEs for Me_n, $n = 2$, 3, 4, to within 0.002 kcal/mol and those for $n = 5$, 6 to better than 0.007 kcal/mol. These errors in BE for the clusters are 1−3 orders of magnitude smaller than those associated with the general dispersion-corrected DFT-based methods, with the largest percent error being ca. 0.4% for the $n = 2$ cluster. This level of performance, combined with the fact that larger clusters will have many of the structural features captured by the clusters with $n = 2-6$, offers us a great deal of confidence in the ability of DCPs to accurately predict the structures and the BEs for larger methane clusters.

We now show how the methane-specific DCPs can be applied to bridge the gap between high-level correlated wave function calculations and molecular mechanics simulations. For this effort, we focus on the minimum energy structures of methane clusters, composed of 2−40 monomers. These systems were recently studied by Takeuchi, who showed that, in general, a Morse potential derived largely from MP2/6-311+G(2df,2pd) calculations on the methane dimer (hereafter the "RP" FF) produced clusters with lower energies than those obtained using the OPLS FF. Takeuchi's work underscores the value of the methodology we propose for bridging CCSD(T)/CBS and FFs to develop a high-quality classical molecular mechanics method: In the search for the most stable methane cluster structures, Takeuchi used a heuristic method combined with geometrical perturbation, with initial structures generated randomly. The number of initial structures was approximately $n \times 100$, with n the number of molecules, with the exception of $n = 3$, for which 10,000 initial structures were necessary. Such searches can be done efficiently by classical molecular dynamics but less-so with DFT-based methods and are virtually impossible for CCSD(T)/CBS.

As described above, we used the PBE0-DCP(methane)/6-31+G(d,p) to compute the forces for each of 35 structures obtained from the stepwise extraction of a single methane molecule from the minimum energy structure of a cluster of eight methane molecules. These calculations can be accomplished in just a few hours using PBE0-DCP(methane)/6-31+G(d,p). The forces generated with the DCP approach were then used to optimize FF parameters associated with eq 3. Surveying the different force-field parameter mixing rules, we ultimately found that the best parameters were obtained using the widely used Lorentz−Berthelot mixing rule (eq 4). For this mixing, a geometrical average is used for the well-depth and an arithmetical average for the collision diameter.[46] The L-J parameters obtained from our efforts are listed in Table 3 and are hereafter referred to as the "TD(Me)" parameters.

Table 3. Optimized Classical Force-Field Lennard-Jones Parameters, Based on the Lorentz−Berthelot Mixing Rule (Eq 4)[a]

atom 1	atom 2	ε	σ
C	C	0.162 204	3.158 188
H	H	0.021 899	2.711 095
C	H	0.059 600	2.934 641

[a]The PBE0-DCP(methane)/6-31+G(d,p) approach was used to generate the force-field to which the parameters were fit (see text for details). The DL_POLY formatted FIELD file used for the optimizations can be found in the Supporting Information

To test the quality of our TD(Me) parameters, we applied them to the methane clusters studies by Takeuchi and compared our cluster binding energies to those he obtained using the OPLS and RP force fields (see Table 1 of ref 36). We did not engage in any conformer searching. Instead, we used both the OPLS and RP structures as starting points for optimizations using our TD(Me) parameters (data referred to as $TD(Me)^{OPLS}$ and $TD(Me)^{RP}$, respectively). To make comparisons meaningful, we performed single-point PBE0-DCP(methane)/6-31+G(d,p) calculations on the two TD(Me) sets of optimized clusters, as well as the cluster sets Takeuchi obtained using the OPLS and RP FFs. The efficiency of the PBE0-DCP approach allows us to more rigorously evaluate the binding energies so obtained for the methane clusters by performing full geometry optimizations with PBE0-DCP-(methane)/6-31+G(d,p) for all of the methane cluster structures obtained using the various FF methods. For a cluster of a given size, the structure that gave the largest binding energy following optimization by PBE0-DCP(methane)/6-31+G(d,p) was used as the reference against which the BEs computed from

dx.doi.org/10.1021/ct4003114 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

FF optimized clusters were compared. This comparison allows for a more direct assessment of the quality of predictions made by the TD(Me) FF parameters against the DCP approach that was used for the development of those FF parameters. The resulting relative binding energies are illustrated graphically in Figure 2, with the BEs tabulated in the Supporting Information.

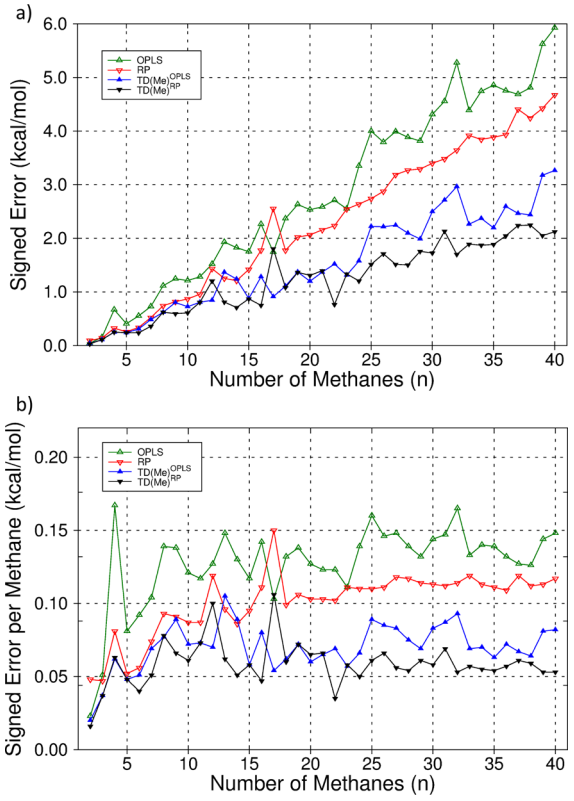

Figure 2. Signed errors in cluster binding energies obtained from various force-field optimization procedures (see text) followed by single-point PBE0-DCP(methane)/6-31+G(d,p) energy calculations. Errors are relative to results obtained from full optimizations performed by PBE0-DCP(methane)/6-31+G(d,p) calculations. The legend indicates the source of the optimized geometries: OPLS and RP force-field optimized structures were taken directly from ref 36; $TD(Me)^{OPLS}$ and $TD(Me)^{RP}$ force-field optimized structures were obtained by starting from the OPLS or RP optimized structures from ref 36 and reoptimizing using the TD(Me) force-field parameters derived in the present work. See text for additional details. (a) Signed error. (b) Signed error per methane. In all cases, optimized monomer structures were used to compute binding energies.

We note that, without exception, clusters that were optimized using PBE0-DCP(methane)/6-31+G(d,p) have BEs that are larger than those obtained from any of the FF optimizations. Consequently, in the foregoing discussions, larger binding energies are synonymous with lower energy (i.e., more stable) clusters.

The data in Figure 2 show that structures obtained using the RP FF have BE values that are in better agreement with the PBE0-DCP(methane) reference values than those obtained with OPLS, with a single clear exception for n = 17. An examination of the structure of this cluster reveals no obvious reason for this species to be an outlier. Nevertheless, our findings do reaffirm the general conclusions of Takeuchi, whose results showed that the RP FF generates lower energy structures for methane clusters than does the OPLS FF. However, the errors in BEs for clusters obtained by both the OPLS and RP optimizations continue to increase with cluster size in a fairly steady fashion: The largest errors in BE are ca. 5.9 (n = 40) kcal/mol for OPLS and 4.7 kcal/mol (n = 40) for RP, values that represent errors of about 8.9% and 6.9%, respectively, relative to the fully optimized PBE0-DCP-(methane), lowest energy structures.

Optimization of the OPLS and RP structures using the TD(Me) FF parameters of this work produces methane clusters with BE values in better agreement with the reference values (i.e., lower energy structures) than those obtained from the OPLS or RP optimizations. Using the better RP structures as starting points allows the TD(Me) FF optimizations to produce clusters with BE values in best agreement with the reference values. Despite the improved performance of the TD(Me) FF parameters for methane cluster optimizations as compared to the OPLS and RP FFs, the errors in predicted BEs are not trivial. For example, the largest errors in BE using $TD(Me)^{OPLS}$ and $TD(Me)^{RP}$ were found to be ca. 4.6 (n = 40) and 2.2 (n = 38) kcal/mol, respectively, or about 40−50% smaller than the errors associated with the OPLS and RP FF optimized clusters. Clearly, and not surprisingly, the four LJ parameters do not possess enough flexibility to fully account for the noncovalent interactions that occur in methane clusters.

In Figure 2b we plot the errors in BE values, predicted using the four FF-based methods, on a per methane basis. The figure shows that the line representing the OPLS data strongly oscillates up to ca. n = 34, whereas the RP errors have appeared to reach a plateau at ca. n = 18. In keeping with the data in Figure 2a, with the RP, errors per methane are generally lower than those obtained with OPLS. A least-squares fitting of the OPLS and RP lines in Figure 2b gives lines with slightly positive slopes (≈0.001 kcal/mol/methane), indicating that there is a slight trend toward growing errors with increasing clusters size. Note that, in the n = 2−40 size range, there is also a trend for binding energy per methane to increase as cluster size increases: At n = 40, BE = 1.66−1.79 kcal/mol/methane, depending on the optimization method utilized. The errors in the OPLS and RP BE/methane data reach values of ca. 0.13 (7.6%) and 0.10 kcal/mol (5.6%) at n = 40, respectively.

The trends in the BE/methane data for the $TD(Me)^{OPLS}$ and $TD(Me)^{RP}$ are somewhat muted compared to those displayed by the OPLS and RP data but display similar oscillatory behavior. Most interesting is that the least-squared fit of the $TD(Me)^{OPLS}$ data gives a horizontal (i.e., zero slope) line with an intercept at about 0.05 kcal/mol/methane, while the $TD(Me)^{RP}$ data gives a line with a slightly negative slope (≈ −0.0003 kcal/mol/methane). At n = 40, the respective errors are less than 0.04 (2.4%) and 0.02 (0.9%) kcal/mol/methane.

Although it can be seen from the data shown in Figure 2 that it is not always possible to unambiguously determine the lowest energy of a methane cluster of a particular size using a single FF method (e.g., clusters with n = 12, 15, 17, 20, etc.), our results demonstrate that the FF parameters derived from a detailed surface generated from PBE0-DCP(methane)/6-31+G(d,p) calculations perform consistently and considerably better than the OPLS and RP FFs. Despite remaining shortcomings, calculations performed with the higher-quality TD(Me) force-field parameters can serve to reduce the initial number of possible low-energy methane cluster structures, which may then be refined by PBE0-DCP(methane) calculations. The necessity

dx.doi.org/10.1021/ct4003114 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

associated with using an approach of this type will obviously increase with cluster size.

CONCLUSIONS

Dispersion-correcting potentials (DCPs) are atom-centered functions that tune the long-range potential in which electrons move. They are used in conjunction with density-functional theory-based methods to improve their ability to describe dispersion and other noncovalent interactions. In this work, we demonstrated how system-specific DCPs could be used to bridge the computational efficiency and accuracy gap that exists between high-level (i.e., complete basis set CCSD(T)) wave function methods and force-field methods. We generated methane-specific DCPs for use with the PBE0/6-31+G(d,p) density-functional. The DCPs significantly reduce the normally large errors associated with the use of PBE0/6-31+G(d,p) to calculate the binding energies of clusters of noncovalently bonded methane molecules. Although we focused our efforts on the PBE0/6-31+G(d,p) approach in this work, we believe that other DFT functionals and basis sets can likewise be employed. In contrast, although the BE values for methane dimers with other dispersion-corrected functionals were fairly good, the quality of the computed BEs for larger methane clusters was inferior. This indicates that the representation of the many-body dispersion interactions plays an important role in dispersion-corrected DFT and requires further study.

The methane-specific DCPs were used to generate high-level (CCSD(T)/CBS-like) data to which Lennard-Jones force-field parameters were fit. We used our DCP approach to compute the forces for each of 35 structures obtained from the stepwise extraction of a single methane molecule from the minimum energy structure of a cluster of eight methane molecules and then used these data to fit methane-specific L-J parameters.

We applied our FF parameters to $(methane)_n$ clusters, with n = 2−40, and compared our calculated binding energies to those obtained by Takeuchi using the OPLS and RP FFs. Our parameters produced BEs that have errors that are more than a factor of 2 lower than those associated with Takeuchi's results on a per methane basis, indicating that the inclusion of accurate dispersion in the reference data has a significant impact on the quality of the FF for clusters structures calculations. Our best results demonstrate that it is possible, using a force-field approach, to obtain binding energies for large methane clusters that have errors below 2%. This result underscores the value of using a density-functional theory-based approach with system-specific dispersion-correcting potentials to bridge the gap between high-level wave function theory and classical force-field methodologies.

ASSOCIATED CONTENT

Supporting Information

Potential energy surfaces computed at various levels of theory for three methane dimers, DL_POLY formatted force field for methane clusters, a Gaussian program input file demonstrating the use of DCPs, binding energies using the various methods described in this work, and the lowest energy PBE0-DCP(methane)/6-31+G(d,p) optimized methane cluster structures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gino.dilabio@nrc.ca.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are grateful to the Centre for Oil Sands Innovation and WestGrid for providing generous support to this work.

REFERENCES

(1) Patchkoskii, S.; Tse, J. S.; Yurchenko, S. N.; Zhechkov, L.; Heine, T.; Seifert, G. Graphene nanostructures as tunable storage media for molecular hydrogen. *Proc. Natl. Acad. Sci. U.S.A* **2005**, *102*, 10439−10444.

(2) Patchkoskii, S.; Tse, J. S. Thermodynamic stability of hydrogen clathrates. *Proc. Natl. Acad. Sci. U.S.A* **2003**, *100*, 14645−14650.

(3) Angell, C. A. Formation of Glasses from Liquids and Biopolymers. *Science* **1995**, *267*, 1924−1935.

(4) 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.

(5) Jorgensen, W. L.; Tirado-Rives, J. The OPLS Force Field for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.

(6) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.

(7) Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V.; Scoles, G. Towards extending the applicability of density functional theory to weakly bound systems. *J. Chem. Phys.* **2001**, *115*, 8748−8757.

(8) Wu, Q.; Yang, W. Empirical correction to density functional theory for van der Waals interactions. *J. Chem. Phys.* **2002**, *116*, 515−524.

(9) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463−1473.

(10) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.

(11) 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.

(12) Johnson, E. R.; Becke, A. D. A post-Hartree−Fock model of intermolecular interactions. *J. Chem. Phys.* **2005**, *123*, 024101.

(13) Becke, A. D.; Arabi, A. A.; Kannemann, F. O. Nonempirical density-functional theory for van der Waals interactions. Can. *J. Chem.* **2010**, *88*, 1057−1062.

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

(15) 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.

(16) Marom, N.; Tkatchenko, A.; Rossi, M.; Gobre, V. V.; Hod, O.; Scheffler, M.; Kronik, L. Dispersion Interactions with Density-Functional Theory: Benchmarking Semiempirical and Interatomic Pairwise Corrected Density Functionals. *J. Chem. Theory Comput.* **2011**, *7*, 3944−3951.

(17) Torres, E.; DiLabio, G. A. A (Nearly) Universally Applicable Method for Modeling Noncovalent Interactions Using B3LYP. *J. Phys. Chem. Lett.* **2012**, *3*, 1738−1744.

(18) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U. Optimization of Effective Atom Centered Potentials for London Dispersion Forces in Density Functional Theory. *Phys. Rev. Lett.* **2004**, *93*, 153004.

(19) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U. Performance of optimized atom-centered potentials for weakly bonded systems using density functional theory. *Phys. Rev. B* **2005**, *71*, 195119.

(20) Lin, I.-C.; Coutinho-Neto, M. D.; Felsenheimer, C.; von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U. Library of dispersion-corrected atom-centered potentials for generalized gradient approximation functionals: Elements H, C, N, O, He, Ne, Ar, and Kr. *Phys. Rev. B* **2007**, *75*, 205131.
(21) For example, Zang et al. use force-fields derived from density-functional theory methods corrected for dispersion using general (and older generation) pair-wise dispersion terms. See: Zang, J.; Nair, S.; Sholl, D. S. *J. Phys. Chem. A* **2013**.
(22) Li, A. H.; Chao, S. D. Intermolecular potentials of the methane dimer calculated with Møller-Plesset perturbation theory and density functional theory. *J. Chem. Phys.* **2006**, *125*, 094312.
(23) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6169.
(24) Christiansen, P. A.; Lee, Y. S.; Pitzer, K. S. Improved ab initio effective core potentials for molecular calculations. *J. Chem. Phys.* **1979**, *71*, 4445.
(25) DCPs were also developed for the silicon atom. See: Johnson, E. R.; DiLabio, G. A. Theoretical Study of Dispersion Binding of Hydrocarbon Molecules to Hydrogen-Terminated Silicon(100)-2 × 1. *J. Phys. Chem. C* **2009**, *113*, 5681−5689.
(26) Mackie, I. D.; DiLabio, G. A. Approximations to complete basis set-extrapolated, highly correlated non-covalent interaction energies. *J. Chem. Phys.* **2011**, *135*, 134318.
(27) The particular CCSD(T)/CBS approach we used modifies the CCSD(T)/aug-cc-pVDZ energy with a correction term obtained by taking the difference between extrapolated MP2/aug-cc-pV(Q-T)Z and MP2/aug-cc-pVDZ energies. All binding energies are obtained by taking the average of binding energies calculated with and without counterpoise corrections.
(28) The structures of the trimer and tetramer were obtained from optimization we performed on structures obtained from reference 36 using a preliminary version of our methane-specific DCPs. Additional information is provided in the Supporting Information Section. Binding energies were computed at various levels of theory 27, including one in which we corrected the CCSD/aug-cc-pVTZ energy with a "triples" (viz., CCSD(T)) correction obtained using aug-cc-pVDZ basis sets and a CBS term obtained from the difference between extrapolated MP2/aug-cc-pV(Q-T)Z and MP2/aug-cc-pVTZ energies. The average of the counterpoise and noncounterpoise corrected binding energies so obtained were used. For additional information, see reference 26. Relaxed monomers were used to obtain binding energies for the clusters.
(29) For example, Tkatchenko and von Lilienfeld pointed out that many DFT methods overestimate many-body interaction terms in noncovalently bonded systems. In principle, DCPs can be designed to overcome this deficiency in much the same way that they are designed to overcome their inability to accurately predict dispersion interactions. This aspect will be examined in more detail in a forthcoming publication.
(30) Tkatchenko, A.; von Lilienfeld, O. A. Popular Kohn-Sham density functionals strongly overestimate many-body interactions in van der Waals systems. *Phys. Rev. B.* **2008**, *78*, 045116.
(31) Gaussian 03, Revision D.01. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.;Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian Inc.: Pittsburgh, PA, 2004.
(32) *Gaussian 09*, Revision C.01; 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, Inc.: Wallingford, CT, 2009.
(33) In refs 31 and 32, the PBE0 functional is referred to as PBE1PBE.
(34) We note that, for hydrocarbons, basis set completeness is relatively easy to achieve and that more extensive basis sets are required to reasonably model systems in which there is a great deal of polarization (e.g., hydrogen bonding).
(35) www.ualberta.ca/~gdilabio.
(36) Takeuchi, H. The structural investigation on small methane clusters described by two different potentials. *Comput. Theor. Chem.* **2012**, *986*, 48−56.
(37) Todorov, I. T.; Smith, W.; Trachenko, K.; Dove, M. T. DL_POLY_3: new dimensions in molecular dynamics simulations via massive parallelism. *J. Mater. Chem.* **2006**, *16*, 1911−1918.
(38) Al-Matar, A. K.; Rockstraw, D. A. A generating equation for mixing rules and two new mixing rules for interatomic potential energy parameters. *J. Comput. Chem.* **2004**, *25*, 660−668.
(39) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.
(40) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom−atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(41) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(42) DiLabio, G. A. Accurate Treatment of van der Waals Interactions using Standard Density Functional Theory with Effective Core-Type Potentials: Application to Carbon-Containing Dimers. *Chem. Phys. Lett.* **2008**, *455*, 348−353.
(43) Mackie, I. D.; DiLabio, G. A. Interactions in Large, Polyaromatic Hydrocarbons Dimers: Application of Density Functional Theory with Dispersion Corrections. *J. Phys. Chem. A* **2008**, *112*, 10968−10976.
(44) Mackie, I. D.; DiLabio, G. A. Accurate dispersion interactions from standard density-functional theory methods with small basis sets. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6092−6098.
(45) Using the damping function described by Becke and Johnson. See ref 12.
(46) For the force-field optimizations, we use a H−C bond length of 1.09354 Å.