JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Evaluating London Dispersion Interactions in DFT: A Nonlocal Anisotropic Buckingham−Hirshfeld Model

A. Krishtal,*[,†,‡] D. Geldof,[†] K. Vanommeslaeghe,[§] C. Van Alsenoy,[†] and P. Geerlings[∥]

[†]Department of Chemistry, University of Antwerp, Universiteitsplein 1, B2610 Antwerp, Belgium

[‡]Fachbereich Chemie, Technische Universität Kaiserslautern, Erwin Schrödinger Straße, D-67663 Kaiserslautern, Germany

[§]Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, 20 Penn St., HSF II-629, Baltimore, Maryland 21201, United States

[∥]Algemene Chemie, Vrije Universiteit Brussel, Pleinlaan 2, B-1050, Brussels, Belgium

ABSTRACT: In this work, we present a novel model, referred to as BH-DFT-D, for the evaluation of London dispersion, with the purpose to correct the performance of local DFT exchange-correlation functionals for the description of van der Waals interactions. The new BH-DFT-D model combines the equations originally derived by Buckingham [Buckingham, A. D. *Adv. Chem. Phys* **1967**, *12*, 107] with the definition of distributed multipole polarizability tensors within the Hirshfeld method [Hirshfeld, F.L. *Theor. Chim. Acta* **1977**, *44*, 129], resulting in nonlocal, fully anisotropic expressions. Since no damping function has been introduced yet into the model, it is suitable in its present form for the evaluation of dispersion interactions in van der Waals dimers with no or negligible overlap. The new method is tested for an extended collection of van der Waals dimers against high-level data, where it is found to reproduce interaction energies at the BH-B3LYP-D/aug-cc-pVTZ level with a mean average error (MAE) of 0.20 kcal/mol. Next, development steps of the model will consist of adding a damping function, analytical gradients, and generalization to a supramolecular system.

1. INTRODUCTION

The problem of description of London dispersion in density functional theory (DFT) using (semi) local exchange-correlation functionals is a well-known problem.[1,2] Since the first diagnostic in 1994,[1] an intense discussion on the origin of the problem and appropriate solutions has been ongoing in the DFT community. Generally, one assumes that the cause lies in the local character of the widely used correlation functionals, which, in contrast to the correlation contribution in post-Hartree−Fock methods such as Møller−Plesset or coupled cluster, only utilize information on the density of the system at one point and are therefore unsuitable for the description of a nonlocal phenomenon such as dispersion. Attempts to introduce nonlocal correlation to DFT, such as the random phase approximation (RPA)[3,4] or the nonlocal van der Waals functionals,[5−8] are being investigated, but unfortunately the improvement comes with a significant increase in the computational cost. Since the relatively low computational cost of DFT is one of the major factors responsible for its status as the most widely used quantum chemical method today, a range of more pragmatic approaches has been developed to correct the performance of DFT for dispersion interactions. Part of these methods rely on reparametrization of existing local correlation functionals,[9−11] motivated by the fact that dispersion is partially included in many functionals and that a suitable reparametrization will allow one to achieve the aspired results more consistently. The drawback of such an approach is that the strong empirical character decreases the reliability. For instance, the performance of the reparametrized functionals often decreases for properties other than the electronic energy. Other attempts are based on adding a correction term, representing the dispersion energy, to the energy calculated using standard DFT methods. Also in this category, one can find highly empirical but computationally attractive methods,[12] based on parameters fitted to reproduce high-level results, as well as the methods with deeper theoretical foundation but computationally more expensive, where *ab initio* information of the systems is used to evaluate the dispersion energy, such as the static or frequency dependent polarizabilities[13−18] or the dipole moment of the exchange-correlation hole (XDM).[19−23] Another noteworthy approach is the adaptation of the symmetry adapted perturbation theory[24] to the framework of DFT, i.e., SAPT(DFT).[25−28] SAPT(DFT) has a significant computational advantage against the highly scaling SAPT as the contribution of intramonomer correlation, already embedded within the Kohn−Sham orbitals, does not need to be evaluated. Although possible to use for the correction of DFT dispersion energies,[29] SAPT(DFT) is mostly meant for an evaluation of the total interaction energy. The explicit expression for the respulsive contribution of electron-exchange to the dispersion energy within SAPT(DFT), though rarely calculated fully due to the computational expense, offers a more theoretically attractive alternative to the empirical damping functions used in other methods. SAPT(DFT) does have the disadvantage of requiring explicit separation of the system in two parts, which makes it impossible for application on intramolecular dispersion interactions, such as those occurring, for example, in biomolecules.

The model described here offers a compromise between computational simplicity and theoretical rigor by combining expressions derived by Buckingham,[30] who used static multipole polarizabilities to express the dispersion interaction between two systems, with the nonempirical Hirshfeld partitioning method. The model offers several advantages: First of all, only *ab initio*

Received: October 11, 2011
Published: November 29, 2011

information about the system is used for the evaluation of the dispersion energy, without making any assumptions about the chemical properties of the system. Second of all, by eliminating frequency-dependence and using *static* atomic polarizabilities, we avoid the necessity to evaluate the computationally expensive Casimir-Polder integrals. Finally, the use of the Hirshfeld method allows us to utilize polarizability *tensors*, which allow one to retain the fully anisotropic character of the dispersion energy.

The model is based on previous work by the present authors,[17,18,23] where a fully anisotropic model was introduced utilizing the Hirshfeld atomic multipole polarizabilities. The effect of anisotropy was isolated by comparing the results with a fully isotropic model and was found to increase the dispersion energies by up to 30%. In this work, the model is extended in two aspects: First of all, we eliminate the pairwise additivity assumption within the dispersion correction, following Nakai and Sato.[14] They found that eliminating the pairwise additivity assumption in their adaptation of Dobson's local-response model[5] of frequency-dependent polarizabilities in the evaluation of dispersion interaction reduced the error on C_6 coefficients to only 6%.[14] Second, we introduce two additional higher correction terms into the model, which now includes terms up to R_{AB}^{-10}. The resulting equations are suitable for the evaluation of dispersion energy in van der Waals dimers, with no or negligible overlap between the densities of the monomers. The next development stages of the model will involve introducing a damping function for shorter intermolecular distances where the effect of exchange becomes important, derivation of analytical gradients, and a generalization to intramolecular dispersion interactions.

The full derivation of the new equations is described in section 2, followed by the computational details in section 3 and test results in section 4. A summary and concluding remarks are given in section 5.

2. METHOD

Within the framework of second-order intermolecular perturbation theory, the London dispersion interaction between two systems A and B is defined as the off-diagonal part of the second-order perturbation energy[31]

$$E_{\mathrm{disp}}^{AB} = -\sum_{j_A \neq n_1}\sum_{j_B \neq n_B} \frac{|\langle \psi_{n_A}^{(0)} \psi_{n_B}^{(0)} | \hat{V}^{AB} | \psi_{j_A}^{(0)} \psi_{j_B}^{(0)} \rangle|^2}{(W_{j_A}^{(0)} - W_{n_A}^{(0)}) + (W_{j_B}^{(0)} - W_{n_B}^{(0)})} \tag{1}$$

where $\psi^{(0)}$ and $W^{(0)}$ are the eigenfunctions and eigenvalues of the unperturbed systems, n the ground states, j the excited states, and $\hat{V}^{AB}$ is the electrostatic Coulombic interaction operator of the two systems:

$$\hat{V}^{AB} = \int\int \frac{\hat{\rho}_A(\mathbf{r})\,\hat{\rho}_B(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \tag{2}$$

In eq 2, $\hat{\rho}(\mathbf{r})$ is the charge density operator, defined as

$$\hat{\rho}(\mathbf{r}) = -\sum_{i=1}^{n_e} \delta(\mathbf{r}-\mathbf{r}_i) + \sum_{j=1}^{N} Z_j\,\delta(\mathbf{r}-\mathbf{R}_j) \tag{3}$$

where the first summation is over the n_e electrons of the system and the second summation is over the N nuclei with charges Z. The total charge density is then given by the expectation value of the charge density operator:

$$\rho(\mathbf{r}) = \langle \psi_n^{(0)} | \hat{\rho}(\mathbf{r}) | \psi_n^{(0)} \rangle \tag{4}$$

This definition of London dispersion assumes a sufficiently large intermolecular distance, such that any electron exchange between the two systems can be neglected.

In order to avoid the pairwise additivity assumption present in our previous model[17,18,23] as well as in the majority of the DFT-D methods,[12,15,19−22] the multiatomic character of the systems can be introduced in the beginning of the derivation.[32] We will do so by considering the molecular charge density operator $\hat{\rho}_A(\mathbf{r})$ at each point of the space $\mathbf{r}$ as a sum of atomic charge density operators:

$$\hat{\rho}_A(\mathbf{r}) = \sum_{a \in A} \hat{\rho}_a(\mathbf{r}) \tag{5}$$

While the division of the nuclear part of the charge density operator between the atoms is trivial, the definition of the electronic part is not unique. However, such a view of molecular density is common within physical-space atom-in-molecule methods that use a fuzzy-atom approach to the definition of atomic density such as the Hirshfeld method[33] or Becke's partitioning method.[2] Using the atomic density operators, the electrostatic Coulomb interaction operator can be rewritten as a pairwise sum of diatomic operators:

$$\hat{V}^{AB} = \sum_{a \in A}\sum_{b \in B} \hat{V}^{ab} \tag{6}$$

where the diatomic operator is defined as

$$\hat{V}^{ab} = \int\int \frac{\hat{\rho}_a(\mathbf{r})\,\hat{\rho}_b(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' = \int \hat{\rho}_a(\mathbf{r})\,\phi_b(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{7}$$

In eq 7, we have defined the atomic potential $\phi_b(\mathbf{r})$ which represents the potential caused at point $\mathbf{r}$ due to the charge distribution of atom b in system B. Following Buckingham's derivation,[30] the atomic potential $\phi_b(\mathbf{r})$ can be expanded in a Taylor series around the origin of atom a, $\mathbf{O}_a$

$$\phi_b(\mathbf{r}) = \phi_b(\mathbf{O}_a) + \sum_i [\nabla_i \phi_b(O_a)](\mathbf{r}-\mathbf{O}_a)_i + \frac{1}{2}\sum_i\sum_j [\nabla_i \nabla_j \phi_b(\mathbf{O}_a)](\mathbf{r}-\mathbf{O}_a)_i(\mathbf{r}-\mathbf{O}_a)_j + \ldots \tag{8}$$

where the subscripts i and j represent the vector components in a Cartesian system of axes: x, y, and z. Subsequently, one can expand the factor $(1)/(|\mathbf{O}_a - \mathbf{r}'|)$ in $\phi_b(\mathbf{O}_a)$ around the origin of atom b, $\mathbf{O}_b$

$$\frac{1}{|\mathbf{O}_a-\mathbf{r}'|} = \frac{1}{|\mathbf{O}_a-\mathbf{O}_b|} + \sum_i \mathbf{T}_{\mathbf{1}i}^{ab}(\mathbf{r}'-\mathbf{O}_b)_i + \frac{1}{2}\sum_i\sum_j T_{\mathbf{2}ij}^{ab}(\mathbf{r}'-\mathbf{O}_b)_i(\mathbf{r}'-\mathbf{O}_b)_j + \ldots \tag{9}$$

where the elements of the $\mathbf{T}_1$ and $\mathbf{T}_2$ tensors represent the first and second derivatives of $R_{ab} = 1/(|\mathbf{O}_a - \mathbf{O}_b|)$ with respect to the components of $\mathbf{R}_{ab}$

$$\mathbf{T}_{\mathbf{1}i}^{ab} = -\frac{\mathbf{R}_{ab,i}}{R_{ab}^3} \tag{10}$$

$$\mathbf{T}_{\mathbf{2}ij}^{ab} = \frac{3\mathbf{R}_{ab,i}\mathbf{R}_{ab,j} - \delta_{ij}R_{ab}^2}{R_{ab}^5} \tag{11}$$

The atomic potential $\phi_b(\mathbf{O}_a)$ can thus be expressed in terms of the atomic charge operators $\hat{q}_b = \int \hat{\rho}(\mathbf{r})\,\mathrm{d}\mathbf{r}$, atomic dipole moment

operators $\hat{\mu}_i^b = \int (\mathbf{r}-\mathbf{O}_b)_i\, \hat{\rho}_b(\mathbf{r})\, d\mathbf{r}$, atomic quadrupole moment operators $\hat{\Theta}_{ij}^b = \int (\mathbf{r}-\mathbf{O}_b)_i (\mathbf{r}-\mathbf{O}_b)_i\, \hat{\rho}_b(\mathbf{r})\, d\mathbf{r}$ and so on.

$$\phi_b(\mathbf{O}_a) = \frac{\hat{q}_b}{R_{ab}} + \sum_i \hat{\mu}_i^b T_{\mathbf{1}i}^{ab} + \frac{1}{2}\sum_{ij} \hat{\Theta}_{ij}^b T_{\mathbf{2}ij}^{ab} + \ldots \quad (12)$$

Inserting eqs 8 and 12 into eq 7 and performing the integration over $\mathbf{r}$ finally results in a diatomic interaction operator expressed in terms of atomic charge and multipole moment operators, and the interatomic distance R_{ab}

$$\hat{V}^{ab} = \frac{\hat{q}_a\hat{q}_b}{R_{ab}} + \sum_i T_{\mathbf{1}i}^{ab}(\hat{q}_a\hat{\mu}_i^b - \hat{\mu}_i^a\hat{q}_b) + \frac{1}{2}\sum_{ij} T_{\mathbf{2}ij}^{ab}(\hat{q}_a\hat{\Theta}_{ij}^b + \hat{\Theta}_{ij}^a\hat{q}_b - \hat{\mu}_i^a\hat{\mu}_j^b) + \ldots \quad (13)$$

By inserting the diatomic interaction operator into eq 1, the total dispersion energy can be expressed as a sum of four-centered atomic contributions

$$E_{\mathrm{disp}}^{AB} = \sum_{a,a'\in A}\sum_{b,b'\in B} E_{\mathrm{disp}}^{aa'bb'} \quad (14)$$

where the first nonzero term in the four-centered atomic dispersion energy is

$$E_{\mathrm{disp}}^{aa'bb'}(R^{-6}) = -\sum_{ijkl} T_{\mathbf{2}ij}^{ab} T_{\mathbf{2}kl}^{a'b'} \sum_{j_A\neq n_A}\sum_{j_B\neq n_B} \times \frac{\langle\psi_{n_A}^{(0)}\psi_{n_B}^{(0)}|\hat{\mu}_i^a\hat{\mu}_j^b|\psi_{j_A}^{(0)}\psi_{j_B}^{(0)}\rangle\langle\psi_{n_A}^{(0)}\psi_{n_B}^{(0)}|\hat{\mu}_k^{a'}\hat{\mu}_l^{b'}|\psi_{j_A}^{(0)}\psi_{j_B}^{(0)}\rangle}{(W_{j_A}^{(0)} - W_{n_A}^{(0)}) + (W_{j_B}^{(0)} - W_{n_B}^{(0)})}$$
$$= -\sum_{ijkl} T_{\mathbf{2}ij}^{ab} T_{\mathbf{2}kl}^{a'b'} \sum_{j_A\neq n_A}\sum_{j_B\neq n_B} \times \frac{\langle\psi_{n_A}^{(0)}|\hat{\mu}_i^a|\psi_{j_A}^{(0)}\rangle\langle\psi_{n_B}^{(0)}|\hat{\mu}_j^b|\psi_{j_B}^{(0)}\rangle\langle\psi_{n_A}^{(0)}|\hat{\mu}_k^{a'}|\psi_{j_A}^{(0)}\rangle\langle\psi_{n_B}^{(0)}|\hat{\mu}_l^{b'}|\psi_{j_B}^{(0)}\rangle}{(W_{j_A}^{(0)} - W_{n_A}^{(0)}) + (W_{j_B}^{(0)} - W_{n_B}^{(0)})} \quad (15)$$

Equation 15 can be connected to the *distributed static atomic polarizabilities* defined, in the framework of perturbation theory, as

$$\alpha_{kl}^{aa'} = 2\sum_{j_A\neq n_A} \frac{\langle\psi_{n_A}^{(0)}|\hat{\mu}_k^a|\psi_{j_A}^{(0)}\rangle\langle\psi_{j_A}^{(0)}|\hat{\mu}_l^{a'}|\psi_{n_A}^{(0)}\rangle}{W_{j_A}^{(0)} - W_{n_A}^{(0)}} \quad (16)$$

This definition of distributed polarizability is connected to the classic sum-over-states expression that makes use of molecular dipole moment operators $\hat{\mu}$:

$$\alpha_{kl} = 2\sum_{j_A\neq n_A} \frac{\langle\psi_{n_A}^{(0)}|\hat{\mu}_k|\psi_{j_A}^{(0)}\rangle\langle\psi_{j_A}^{(0)}|\hat{\mu}_l|\psi_{n_A}^{(0)}\rangle}{W_{j_A}^{(0)} - W_{n_A}^{(0)}} \quad (17)$$

Equation 17 is equivalent to the coupled perturbed Hartree−Fock expression

$$\alpha_{kl} = -tr[\mathbf{H}^{(k)}\mathbf{D}^{(l)}] = -\int \mathbf{r}_k\, \rho^{(l)}(\mathbf{r})\, d\mathbf{r} \quad (18)$$

where $\mathbf{H}^{(k)}$ is the dipole moment matrix, $\mathbf{D}^{(l)}$ is the perturbed density matrix, and $\rho^{(l)}(\mathbf{r})$ is the perturbed molecular density.

In order to connect eq 15 to eq 16, the sum in the denominator in eq 15 is replaced, as suggested by Buckingham,[30] by a product, by introducing average excitation energies U:

$$\frac{1}{(W_{j_A}^{(0)} - W_{n_A}^{(0)}) + (W_{j_B}^{(0)} - W_{n_B}^{(0)})} = \frac{U_A U_B(1+\Delta)}{(U_A+U_B)(W_{j_A}^{(0)} - W_{n_A}^{(0)})(W_{j_B}^{(0)} - W_{n_B}^{(0)})} \quad (19)$$

and neglecting Δ which is given by

$$\Delta = \frac{[U_A^{-1} + U_B^{-1}] - [(W_{j_A}^{(0)} - W_{n_A}^{(0)}) + (W_{j_B}^{(0)} - W_{n_B}^{(0)})]}{(W_{j_A}^{(0)} - W_{n_A}^{(0)}) + (W_{j_B}^{(0)} - W_{n_B}^{(0)})} \quad (20)$$

As one can see, the error made by the approximation is minimal when the average excitation energies U are similar to the energies W of the lowest excited states.

The elements of the atomic static distributed polarizability tensors $\alpha_{ij}^{aa'}$ can be obtained by means of the iterative Hirshfeld method.[33,34] In the Hirshfeld method, the atomic region is defined through a weight function constructed from the promolecular atomic densities $\rho_a^{\mathrm{pro}}(\mathbf{r})$

$$w_a(r) = \frac{\rho_a^{\mathrm{pro}}(\mathbf{r})}{\sum_{a'\in A}\rho_{a'}^{\mathrm{pro}}(\mathbf{r})} \quad (21)$$

The fuzzy atom model of the Hirshfeld method is therefore in perfect agreement with the expression of the molecular charge density operators in terms of atomic charge density operators introduced at the beginning of the derivation in eq 5. In the iterative Hirshfeld method (H−I), the spherically symmetric promolecular atomic densities, which represent the densities of the (fictive) atoms prior to bonding, are optimized self-consistently to satisfy the constraint that the number of electrons in a promolecular atom is identical to the number of electrons in the atom in the molecule, i.e.:

$$N_a^{\mathrm{pro}} = \int \rho_a^{\mathrm{pro}}(\mathbf{r})\, d\mathbf{r} = \int w_a(\mathbf{r})\,\rho(\mathbf{r})\, d\mathbf{r} = N_a \quad (22)$$

As a result, the H−I weight function represents the best definition of an atom in molecule, as retrieved through information theory.[35] In particular, the atomic region within the H−I method is defined without any use of parameters or any prior knowledge of the system and therefore reflects the specific chemical surroundings of the atom.

Previously, static atomic intrinsic polarizabilities have been defined within the Hirshfeld method[36] by Kristhal et al. and used in our previous pairwise additive model for the evaluation of dispersion interactions.[17] We extend here the definition further to *distributed* atomic intrinsic polarizabilities by partitioning the dipole response μ_k to the applied dipole field μ_j between two different atoms a and a'

$$\alpha_{ij}^{aa'} = -\int (\mathbf{r}-\mathbf{O}_a)_i\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(j)_{a'}}(\mathbf{r})\, d\mathbf{r} \quad (23)$$

The partitioning is thus realized by introducing two weight functions into the definition of polarizability in eq 18 and centering the quantities on the origins of the atoms. The use of two weight functions for the partitioning of a property between two atoms has been previously introduced in the framework of the Hirshfeld method for the partitioning of overlap populations,[37] MP2 correlation energy,[38] and the definition of

atomic density matrices.[39,40] In eq 23, $\rho^{(j)_{a'}}(\mathbf{r})$ stands for the molecular density perturbed by a dipole field centered on the atom a'. For systems with zero net charge, $\rho^{(j)_{a'}}(\mathbf{r}) = \rho^{(j)}(\mathbf{r})$ due to the origin independence of the dipole moment in neutral systems. However, for higher multipole moments, the perturbed density requires explicit translation to the center of the atom.[42] Note that our definition of distributed atomic polarizability differs from previous definitions utilized in the context of evaluation of dispersion interactions. For instance, Hättig et al.[41] have used Stone's model for the calculation of topologically partitioned polarizabilities in a multicenter multipole expansion,[43] obtained from a partitioning of the molecular volume according to Bader's QTAIM method.[44] Sato and Nakai,[13] on the other hand, have developed a distributed frequency dependent polarizability model based on Dobson's[5] local-approximation model.

The elements of the dipole−quadrupole distributed atomic intrinsic polarizability tensor $\mathbf{A}$ can be partitioned in two ways. On the one hand, one can consider the dipole response μ_i to the applied quadrupole field Θ_{jk}

$$\mathbf{A}_1{}^{aa'}_{i,jk} = -\int (\mathbf{r}-\mathbf{O}_a)_i\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(jk)_{a'}}(\mathbf{r})\, d\mathbf{r} \tag{24}$$

On the other hand, one can consider the quadrupole response Θ_{ij} to the applied dipole field μ_k

$$\mathbf{A}_2{}^{aa'}_{ij,k} = -\int (\mathbf{r}-\mathbf{O}_a)_i(\mathbf{r}-\mathbf{O}_a)_j\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(k)_{a'}}(\mathbf{r})\, d\mathbf{r} \tag{25}$$

Due to the origin dependence of the quadrupole moment, the quadrupole field perturbed density in eq 24 is centered on atom a' by translating it as follows:[42]

$$\rho^{(ij)_{a'}}(\mathbf{r}) = \rho^{(ij)}(\mathbf{r}) - a_i'\rho^{(j)}(\mathbf{r}) - a_j'\rho^{(i)}(\mathbf{r}) \tag{26}$$

The results obtained from the two definitions are not equivalent, and both are integrated within the present dispersion model by partitioning systematically the response on atoms a (or b) and the applied field on atoms a' (or b'). Since the summation in eq 14 is for both a and a' (or b and b') over all atoms in system A (or B), both possibilities are included for each combination of atoms a and a' (or b and b'). Note that following the notation in eqs 24 and 25, $\mathbf{A_1}$ is a 3×9 tensor and $\mathbf{A_2}$ is a 9×3 tensor. The quadrupole ($\mathbf{C}$), octupole ($\mathbf{R}$), dipole-octupole ($\mathbf{E_1}$ and $\mathbf{E_2}$), and quadrupole-octupole ($\mathbf{H_1}$ and $\mathbf{H_2}$) polarizabilities used in this work are defined in a similar fashion, as given in eqs 27−32.

$$\mathbf{C}^{aa'}_{ij,kl} = -\int (\mathbf{r}-\mathbf{O}_a)_i(\mathbf{r}-\mathbf{O}_a)_j\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(kl)_{a'}}(\mathbf{r})\, d\mathbf{r} \tag{27}$$

$$\mathbf{R}^{aa'}_{ijk,lmn} = -\int (\mathbf{r}-\mathbf{O}_a)_i(\mathbf{r}-\mathbf{O}_a)_j(\mathbf{r}-\mathbf{O}_a)_k\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(lmn)_{a'}}(\mathbf{r})\, d\mathbf{r} \tag{28}$$

$$\mathbf{E}_1{}^{aa'}_{i,jkl} = -\int (\mathbf{r}-\mathbf{O}_a)_i\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(jkl)_{a'}}(\mathbf{r})\, d\mathbf{r} \tag{29}$$

$$\mathbf{E}_2{}^{aa'}_{ijk,l} = -\int (\mathbf{r}-\mathbf{O}_a)_i(\mathbf{r}-\mathbf{O}_a)_j(\mathbf{r}-\mathbf{O}_a)_k\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(l)_{a'}}(\mathbf{r})\, d\mathbf{r} \tag{30}$$

$$\mathbf{H}_1{}^{aa'}_{ij,klm} = -\int (\mathbf{r}-\mathbf{O}_a)_i(\mathbf{r}-\mathbf{O}_a)_j\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(klm)_{a'}}(\mathbf{r})\, d\mathbf{r} \tag{31}$$

$$\mathbf{H}_2{}^{aa'}_{ijk,lm} = -\int (\mathbf{r}-\mathbf{O}_a)_i(\mathbf{r}-\mathbf{O}_a)_j(\mathbf{r}-\mathbf{O}_a)_k\, w_a(\mathbf{r})\, w_{a'}(\mathbf{r})\, \rho^{(lm)_{a'}}(\mathbf{r})\, d\mathbf{r} \tag{32}$$

In eqs 28, 29, and 31, $\rho^{(ijk)_{a'}}(\mathbf{r})$ is translated to the center of atom a' by

$$\rho^{(ijk)_{a'}}(\mathbf{r}) = \rho^{(ijk)}(\mathbf{r}) - a_i'\rho^{(jk)}(\mathbf{r}) - a_j'\rho^{(ik)}(\mathbf{r}) - a_k'\rho^{(ij)}(\mathbf{r}) + a_i'a_j'\rho^{(k)}(\mathbf{r}) + a_i'a_k'\rho^{(j)}(\mathbf{r}) + a_j'a_k'\rho^{(i)}(\mathbf{r}) \tag{33}$$

The first term in the four-centered atomic dispersion energy in eq 15 can be expressed in terms of the Hirshfeld distributed atomic polarizabilities

$$E_{\mathrm{disp}}^{aa'bb'}(R^{-6}) = -\frac{U_AU_B}{U_A+U_B}\sum_{ijkl} T_{ij}^{ab}T_{kl}^{a'b'}\alpha_{ik}^{aa'}\alpha_{jl}^{bb'} = -\frac{U_AU_B}{U_A+U_B}\, tr[\mathbf{T_2}^{ab}\boldsymbol{\alpha}^{bb'}\mathbf{T_2}^{a'b'}\boldsymbol{\alpha}^{aa'\tau}] \tag{34}$$

where $\mathbf{T}_2$ and α are 3×3 tensors and τ designates a transposed matrix. Note that the standard R^{-6} dependence of the first term only becomes explicit when $a = a'$ and $b = b'$. However, in the long-range distance, where the intermolecular distance R_{AB} is significantly larger than the interatomic distances $R_{aa'}$ and $R_{bb'}$, the correct R^{-6} asymptotic behavior of the dispersion energy is reproduced. The higher terms of the dispersion energy in this model are given by

$$E_{\mathrm{disp}}^{aa'bb'}(R^{-7}) = -\frac{U_AU_B}{U_A+U_B}\left\{\frac{1}{2}tr[\mathbf{T_2}^{ab}\mathbf{A_2}^{bb'}\mathbf{T_3}^{a'b'\tau}\boldsymbol{\alpha}^{aa'\tau}] - \frac{1}{2}tr[\mathbf{T_2}^{ab}\boldsymbol{\alpha}^{bb'}\mathbf{T_3}^{a'b'}\mathbf{A_1}^{aa'\tau}] + \frac{1}{2}tr[\mathbf{T_3}^{ab}\mathbf{A_2}^{bb'}\mathbf{T_2}^{a'b'}\boldsymbol{\alpha}^{aa'\tau}] - \frac{1}{2}tr[\mathbf{T_3}^{ab}\mathbf{A_2}^{aa'}\mathbf{T_2}^{a'b'}\boldsymbol{\alpha}^{bb'\tau}]\right\} \tag{35}$$

$$E_{\mathrm{disp}}^{aa'bb'}(R^{-8}) = -\frac{U_AU_B}{U_a+U_b}\left\{\frac{1}{4}tr[\mathbf{T_3}^{ab}\mathbf{C}^{bb'}\mathbf{T_3}^{a'b'\tau}\boldsymbol{\alpha}^{aa'\tau}] + \frac{1}{4}\mathrm{tr}[\mathbf{T_3}^{ab}\mathbf{C}^{aa'}\mathbf{T_3}^{a'b'\tau}\boldsymbol{\alpha}^{bb'\tau}] - \frac{1}{4}tr[\mathbf{T_2}^{ab}\mathbf{A_2}^{bb'}\mathbf{T_{4(2)}}^{a'b'}\mathbf{A_1}^{aa'\tau}] - \frac{1}{4}tr[\mathbf{T_3}^{ab}\mathbf{A_1}^{bb'}\mathbf{T_3}^{a'b'}\mathbf{A_1}^{aa'\tau}] - \frac{1}{4}tr[\mathbf{T_3}^{ab}\mathbf{A_2}^{aa'}\mathbf{T_3}^{a'b'}\mathbf{A_1}^{bb'\tau}] - \frac{1}{4}tr[\mathbf{T_{4(2)}}^{ab}\mathbf{A_2}^{bb'}\mathbf{T_2}^{a'b'}\mathbf{A_2}^{aa'\tau}] + \frac{1}{6}tr[\mathbf{T_2}^{ab}\mathbf{E_1}^{bb'}\mathbf{T_{4(1)}}^{a'b'\tau}\boldsymbol{\alpha}^{aa'\tau}] + \frac{1}{6}tr[\mathbf{T_2}^{ab}\boldsymbol{\alpha}^{bb'}\mathbf{T_{4(1)}}^{a'b'}\mathbf{E_1}^{aa'\tau}] + \frac{1}{6}tr[\mathbf{T_{4(1)}}^{ab}\mathbf{E_2}^{bb'}\mathbf{T_2}^{a'b'}\boldsymbol{\alpha}^{aa'\tau}] + \frac{1}{6}tr[\mathbf{T_{4(1)}}^{ab}\mathbf{E_2}^{aa'}\mathbf{T_2}^{a'b'}\boldsymbol{\alpha}^{bb'\tau}]\right\} \tag{36}$$

$$E_{\mathrm{disp}}^{aa'bb'}(R^{-9}) = -\frac{U_AU_B}{U_A+U_B}\left\{\frac{1}{12}tr[\mathbf{T_2}^{ab}\mathbf{A_1}^{bb'}\mathbf{T_5}^{a'b'}\mathbf{E_1}^{aa'\tau}] - \frac{1}{12}tr[\mathbf{T_2}^{ab}\mathbf{E_1}^{bb'}\mathbf{T_5}^{a'b'}\mathbf{A_1}^{aa'\tau}] + \frac{1}{12}tr[\mathbf{T_3}^{ab}\mathbf{H_1}^{bb'}\mathbf{T_{4(1)}}^{a'b'\tau}\boldsymbol{\alpha}^{aa'\tau}]\right.$$

$$-\frac{1}{12}tr[\mathbf{T}_3{}^{ab}\mathbf{H}_1{}^{aa'}\mathbf{T}_{4(1)}{}^{a'b'\tau}\boldsymbol{\alpha}^{bb'\tau}] + \frac{1}{8}tr[\mathbf{T}_3{}^{ab}\mathbf{C}^{aa'}\mathbf{T}_{4(2)}{}^{a'b'}\mathbf{A}_1{}^{bb'\tau}]$$
$$-\frac{1}{8}tr[\mathbf{T}_3{}^{ab}\mathbf{C}^{bb'}\mathbf{T}_{4(2)}{}^{a'b'}\mathbf{A}_1{}^{aa'\tau}] + \frac{1}{12}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{H}_2{}^{bb'}\mathbf{T}_3{}^{a'b'\tau}\boldsymbol{\alpha}^{aa'\tau}]$$
$$-\frac{1}{12}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{H}_2{}^{aa'}\mathbf{T}_3{}^{a'b'\tau}\boldsymbol{\alpha}^{bb'\tau}] + \frac{1}{12}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{E}_2{}^{aa'}\mathbf{T}_3{}^{a'b'}\mathbf{A}_1{}^{bb'\tau}]$$
$$-\frac{1}{12}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{E}_2{}^{bb'}\mathbf{T}_3{}^{a'b'}\mathbf{A}_1{}^{aa'\tau}] + \frac{1}{8}tr[\mathbf{T}_{4(2)}{}^{ab}\mathbf{A}_2{}^{bb'}\mathbf{T}_3{}^{a'b'}\mathbf{C}^{aa'\tau}]$$
$$-\frac{1}{8}tr[\mathbf{T}_{4(2)}{}^{ab}\mathbf{C}^{bb'}\mathbf{T}_3{}^{a'b'\tau}\mathbf{A}_2{}^{aa'\tau}] + \frac{1}{12}tr[\mathbf{T}_5{}^{ab}\mathbf{E}_2{}^{aa'}\mathbf{T}_2{}^{a'b'}\mathbf{A}_2{}^{bb'\tau}]$$
$$-\frac{1}{12}tr[\mathbf{T}_5{}^{ab}\mathbf{E}_2{}^{bb'}\mathbf{T}_2{}^{a'b'}\mathbf{A}_2{}^{aa'\tau}]\Big\} \quad (37)$$

$$E_{\mathrm{disp}}^{aa'bb'}(R^{-10}) = -\frac{U_A U_B}{U_A + U_B}\Big\{\frac{1}{36}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{R}^{bb'}\mathbf{T}_{4(1)}{}^{a'b'\tau}\boldsymbol{\alpha}^{aa'\tau}]$$
$$+\frac{1}{36}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{R}^{aa'}\mathbf{T}_{4(1)}{}^{a'b'\tau}\boldsymbol{\alpha}^{bb'\tau}] + \frac{1}{36}tr[\mathbf{T}_2{}^{ab}\mathbf{E}_1{}^{bb'}\mathbf{T}_6{}^{a'b'}\mathbf{E}_1{}^{aa'\tau}]$$
$$+\frac{1}{36}tr[\mathbf{T}_6{}^{ab}\mathbf{E}_2{}^{bb'}\mathbf{T}_2{}^{a'b'}\mathbf{E}_2{}^{aa'\tau}] + \frac{1}{36}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{E}_2{}^{bb'}\mathbf{T}_{4(1)}{}^{a'b'}\mathbf{E}_1{}^{aa'\tau}]$$
$$+\frac{1}{36}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{E}_2{}^{aa'}\mathbf{T}_{4(1)}{}^{a'b'}\mathbf{E}_1{}^{aa'\tau}] + \frac{1}{24}tr[\mathbf{T}_3{}^{ab}\mathbf{C}^{bb'}\mathbf{T}_5{}^{a'b'}\mathbf{E}_1{}^{aa'\tau}]$$
$$+\frac{1}{24}tr[\mathbf{T}_3{}^{ab}\mathbf{C}^{aa'}\mathbf{T}_5{}^{a'b'}\mathbf{E}_1{}^{bb'\tau}] + \frac{1}{24}tr[\mathbf{T}_5{}^{ab}\mathbf{E}_2{}^{bb'}\mathbf{T}_3{}^{a'b'}\mathbf{C}^{aa'\tau}]$$
$$+\frac{1}{24}tr[\mathbf{T}_5{}^{ab}\mathbf{E}_2{}^{aa'}\mathbf{T}_3{}^{a'b'}\mathbf{C}^{bb'\tau}] - \frac{1}{24}\mathrm{tr}[\mathbf{T}_3{}^{ab}\mathbf{H}_1{}^{bb'}\mathbf{T}_5{}^{a'b'\tau}\mathbf{A}_1{}^{aa'\tau}]$$
$$-\frac{1}{24}tr[\mathbf{T}_3{}^{ab}\mathbf{H}_1{}^{aa'}\mathbf{T}_5{}^{a'b'\tau}\mathbf{A}_1{}^{bb'\tau}] - \frac{1}{24}tr[\mathbf{T}_5{}^{ab}\mathbf{H}_2{}^{bb'}\mathbf{T}_3{}^{a'b'\tau}\mathbf{A}_2{}^{aa'\tau}]$$
$$-\frac{1}{24}tr[\mathbf{T}_5{}^{ab}\mathbf{H}_2{}^{aa'}\mathbf{T}_3{}^{a'b'\tau}\mathbf{A}_2{}^{bb'\tau}] - \frac{1}{24}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{H}_2{}^{bb'}\mathbf{T}_{4(2)}{}^{a'b'\tau}\mathbf{A}_1{}^{aa'\tau}]$$
$$-\frac{1}{24}tr[\mathbf{T}_{4(1)}{}^{ab}\mathbf{H}_2{}^{aa'}\mathbf{T}_{4(2)}{}^{a'b'\tau}\mathbf{A}_1{}^{bb'\tau}] - \frac{1}{24}tr[\mathbf{T}_{4(2)}{}^{ab}\mathbf{H}_1{}^{bb'}\mathbf{T}_{4(1)}{}^{a'b'\tau}\mathbf{A}_2{}^{aa'\tau}]$$
$$-\frac{1}{24}tr[\mathbf{T}_{4(2)}{}^{ab}\mathbf{A}_2{}^{bb'}\mathbf{T}_{4(1)}{}^{a'b'\tau}\mathbf{H}_1{}^{aa'\tau}] + \frac{1}{16}tr[\mathbf{T}_{4(2)}{}^{ab}\mathbf{C}^{bb'}\mathbf{T}_{4(2)}{}^{a'b'}\mathbf{C}^{aa'\tau}]\Big\} \quad (38)$$

In eqs 35−38, $\mathbf{T}_3$ is a 3×9 tensor, $\mathbf{T}_{4(1)}$ is a 3×27 tensor, $\mathbf{T}_{4(2)}$ is a 9×9 tensor, $\mathbf{T}_5$ is a 9×27 tensor, and $\mathbf{T}_6$ is a 27×27 tensor. The use of the tensors, as opposed to isotropic values of the polarizabilities, allows one to retain the full anisotropic character of the equations. The effect of anisotropy was previously shown by us to contribute as much as 30% to the dispersion energy;[17,18] neglecting of the anisotropy results in a loss of all of the terms involving **A**, **E**, and **H** and also results in lower values of terms involving α, **C**, and **R**.

The average excitation energies U_A and U_B are approximated by

$$U_A = \frac{2}{3}\frac{\sum_{a \in A}\langle \mu_a^2 \rangle_{\mathrm{XDM}}}{\sum_{a \in A}\alpha_a^{\mathrm{iso}}} \quad (39)$$

where $\langle \mu_a{}^2 \rangle_{\mathrm{XDM}}$ is the expectation value of the square of the atomic exchange-hole dipole moment[19] and $\alpha_a{}^{\mathrm{iso}}$ is the isotropic value of the (local) atomic polarizability,[36] both obtained using the iterative Hirshfeld method.[34] Note that in our previous pairwise additive model[17,18,23] the excitation energies were atomic properties, located inside the summation over the atoms. This difference prevents us from making a direct comparison with previously obtained results in order to evaluate the effect of the elimination of the pairwise additivity. The downside of the elimination of the pairwise additivity is the formal $\mathcal{O}(N^4)$ scaling as opposed to the $\mathcal{O}(N^2)$ scaling in the pairwise additive model. However, the steep R^{-n} dependence of the dispersion energy expression can be exploited in order to reduce the scaling behavior to $\mathcal{O}(N^2)$ in larger systems.

Finally, it should be noted that while in the derivation presented here we have used the primitive form of the quadrupole and octupole moments, an analogue derivation can be made using the traceless definition of these properties, leading to expressions similar to eqs 34−38 and identical numerical results for the dispersion energy.

3. COMPUTATIONAL DETAILS

The geometries of the dimers considered in this work were taken from the S22[45] and SCAI[46] data sets as well as from ref 18, where all high level interaction energies were obtained at the CCSD(T)/CBS level. The evaluation of dispersion energies is done for the B3LYP functional using the aug-cc-pVTZ and 6-311++G(2df,p) basis sets. This is achieved in three steps: In the first step, the molecular multipole polarizabilities of the monomers are calculated by performing single point calculations with the multipole fields applied in the positive and negative directions with the strength of 0.0001 au using the B3LYP/aug-cc-pVTZ and B3LYP/6-311++G(2df,p) functional and basis set combinations in the Gaussian 09 program.[47] Consequently, the first order perturbed density matrices are obtained using the finite field method in the BRABO program.[49] In the second step, the multipole polarizabilities are partitioned into atomic contributions using the STOCK program.[48] In the last step, the dispersion energies are calculated using the ATDISP program.[23] The interaction energies at the DFT level are obtained using Gaussian 09 with the B3LYP/aug-cc-pVTZ and B3LYP/6-311++G(2df,p) functional and basis set combinations, utilizing the counterpoise method[50] for correction of the basis set superposition error.

4. RESULTS AND DISCUSSION

As an initial test of the performance of the proposed model, the interaction energies of 34 van der Waals dimers are compared with data obtained at the CCSD(T)/CBS level of theory. The dimers in the test set are gathered from three different databases, namely the S22 database for benchmark on noncovalent complexes,[45] the SCAI data set containing amino acid side chain interactions,[46] and several complexes optimized by the present authors in previous work.[18] Since the present model does not take into account the repulsive effect of exchange at shorter distances and no damping function has been implemented at this point, only dimers with sufficiently large interatomic distances are considered, i.e., only dispersion and/or induction bonded dimers. The evaluation of the performance of the model for systems stabilized by shorter-ranged interactions such as hydrogen bonds is postponed to a later stage.

The total interaction energy is obtained by supplementing the BSSE-corrected DFT interaction energy by the dispersion energy obtained using the BH-DFT-D model. Whereas the electrostatic and induction interactions between the monomers are generally assumed to be modeled correctly by (semi)-local exchange functionals, the performance of the functionals for the description of dispersion interactions varies substantially. In order to prevent double-counting, one wishes to utilize a functional which is quasi "dispersion free". In our previous work,[18] we have observed that the B3LYP functional consequently predicts repulsive interaction energies for pure dispersion-bonded complexes, while, for example the PBE functional

Table 1. The Interaction Energies of a Set of van der Waals Dimers[a]

		aug-cc-pVTZ		6-311++G(2df,p)	
dimer	CCSD(T)/CBS	B3LYP	BH-B3LYP-D	B3LYP	BH-B3LYP-D
He_2[(a)]	−0.02	0.04	−0.00	0.05	0.03
He−Ne[(a)]	−0.04	0.04	−0.02	0.04	0.02
He−Ar[(a)]	−0.06	0.07	−0.02	0.08	0.03
Ne_2[(a)]	−0.07	0.05	−0.05	0.03	−0.03
Ne−Ar[(a)]	−0.12	0.02	−0.14	0.07	−0.03
Ar_2[(a)]	−0.27	0.17	−0.19	0.17	−0.08
He−N_2 L-shaped[(a)]	−0.04	0.09	−0.01	0.10	0.05
He−N_2 T-shaped[(a)]	−0.06	0.11	−0.03	0.12	0.06
He−FCl[(a)]	−0.10	0.07	−0.07	0.08	0.01
FCl−He[(a)]	−0.13	0.05	−0.34	0.06	−0.16
Ne−CH_4[(a)]	−0.18	0.16	−0.20	0.15	−0.07
CH_4−C_2H_4[(a)]	−0.45	0.40	−0.56	0.39	−0.33
SiH_4−CH_4[(a)]	−0.82	0.55	−1.03	0.59	−0.47
$(OCS)_2$[(a)]	−1.76	0.76	−1.71	0.74	−1.05
C_6H_6−CH_4[(b)]	−1.50	0.77	−1.72	0.76	−1.12
CH_4−CH_4[(b)]	−0.53	0.38	−0.49	0.40	−0.19
$(C_2H_4)_2$[(b)]	−1.51	0.49	−1.68	0.51	−0.87
$(C_4H_4N_2)_2$[(b)]	−4.42	2.44	−4.52	2.49	−3.83
C_2H_4−C_2H_2[(b)]	−1.53	−0.66	−1.94	−0.64	−1.70
C_6H_6−H_2O[(b)]	−3.28	−1.20	−4.03	−1.36	−3.49
C_6H_6−NH_3[(b)]	−2.35	−0.11	−2.67	−0.18	−2.16
C_6H_6−HCN[(b)]	−4.46	−1.98	−4.86	−2.06	−4.23
$(C_6H_6)_2$ parallel-displaced[(b)]	−2.73	3.70	−3.59	3.75	−3.07
$(C_6H_6)_2$ T-shaped[(b)]	−2.74	0.98	−2.64	0.96	−1.89
C_6H_6−C_8H_7N T-shaped[(b)]	−5.73	−0.55	−5.55	−0.54	−4.61
Ala−Leu[(c)]	−1.07	0.84	−1.21	0.88	−0.66
Ile−Ile[(c)]	−1.24	0.72	−1.21	0.75	−0.75
Ile−Leu[(c)]	−1.39	0.41	−0.90	0.43	−0.65
Leu−Gly[(c)]	−0.77	0.26	−0.48	0.27	−0.27
Leu−Leu[(c)]	−1.62	0.40	−0.93	0.43	−0.72
Leu−Thr[(c)]	−1.09	0.37	−0.85	0.39	−0.60
Met−Met[(c)]	−2.03	1.69	−2.36	1.74	−1.47
Val−Leu[(c)]	−1.08	0.40	−0.84	0.42	−0.57
Val−Val[(c)]	−1.39	0.75	−1.41	0.80	−1.03
MAE			0.20		0.36
MAX			0.86		−1.12
R			0.98		0.97

[a] The side chains of the amino acids are noted by the standard three letter codes. The geometries and reference data are obtained from refs 18[(a)], 45[(b)], and 46[(c)]. The dispersion corrected values are calculated using eqs 34−38. All interaction energy values are given in kcal/mol. MAE is the unsigned mean absolute error. MAX is the signed maximal absolute error, and *R* is the correlation coefficient.

already recovers a part of the dispersion energy. For this reason, all "pure" DFT interaction energies in this work are obtained using the B3LYP functional. In ref 17, we have also shown that the dispersion energy is not sensitive to the nature of the functional: very similar dispersion energies were obtained using B3LYP, PBE, and TPSS functionals. Although the precise equations of the model are different in the current work due to, among other things, elimination of the pairwise additivity, we expect the robustness of the model with respect to the choice of the exchange-correlation functional to remain since it mainly depends on the reproduction of static multipole polarizability values. On the other hand, the effect of the size of the basis set utilized for the calculation of the pure DFT interaction energies and the multipole polarizabilities is investigated by examining a large (aug-cc-pVTZ) and a medium-sized (6-311++G(2df,p)) basis set.

The interaction energies of the selected dimers are listed in Table 1. The difference between the pure B3LYP interaction energies calculated using the aug-cc-pVTZ and the 6-311++G-(2df,p) basis sets is mostly small, with a mean absolute error of 0.03 kcal/mol. The interaction energies obtained with the smaller basis set are mostly slightly higher. The largest deviations are observed for the three induction-bonded dimers C_6H_6−H_2O, C_6H_6−NH_3, and C_6H_6−HCN where the B3LYP interaction energy is already attractive: the values obtained using the 6-311+

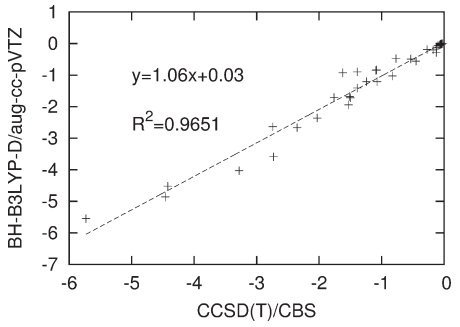

Figure 1. The linear regression parameters for the BH-B3LYP-D/aug-cc-pVTZ model. All values are in kcal/mol.

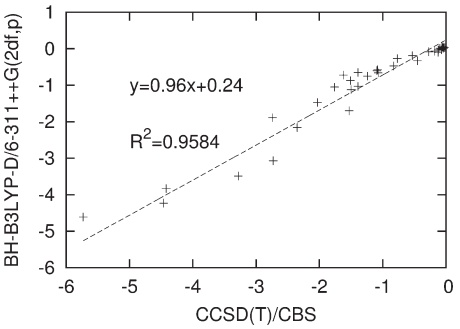

Figure 2. The linear regression parameters for the BH-B3LYP-D/6-311++G(2df,p) model. All values are in kcal/mol.

+G(2df,p) basis set are lower by −0.17 kcal/mol for the C_6H_6−H_2O dimer and −0.08 kcal/mol for the C_6H_6−NH_3 and C_6H_6−HCN dimers.

In contrast to the B3LYP interaction energies, the dispersion energies obtained using the two basis sets vary significantly and, as a result, also the total BH-B3LYP-D interaction energy values. The dispersion energies obtained using the smaller basis set are considerably smaller, underestimated by an average 30.2%. This is reflected in the (unsigned) mean absolute error (MAE) and (signed) maximum absolute error (MAX) listed in Table 1 and the linear regression coefficients shown in Figure 1 for the aug-cc-pVTZ basis set and in Figure 2 for the 6-311++G(2df,p) basis set. Despite the underestimation of the dispersion energy values, the regression coefficient for the 6-311++G(2df,p) is very high (R = 0.97). The MAE and MAX values, on the other hand, are considerably higher for 6-311++G(2df,p) than for the larger basis set. It should be noted that while the largest error for the 6-311++G(2df,p) basis set is for the C_6H_6−C_8H_7N T-shaped dimer, where the BH-B3LYP-D interaction energy is underestimated by −1.12 kcal/mol compared to the CCSD(T)/CBS value, the largest error for BH-D3LYP/aug-cc-pVTZ is an overestimation of the interaction energy of the $(C_6H_6)_2$ parallel-displaced dimer. The origin for this overestimation lies in the significant contribution of the exchange at the van der Waals minimum: this is reflected in the relative values of the disperion energy series for this system, where $E_{disp}(R^{-6}) \approx E_{disp}(R^{-8})$ (cf. Figure 3 below). This indicates that the series crossed the point of convergence and the exchange-uncorrected multipole expansion becomes invalid. A similar situation is encountered for the induction-bonded C_2H_4−C_2H_2 dimer. Other overestimation errors are found in induction-bonded dimers C_6H_6−H_2O, C_6H_6−NH_3, and C_6H_6−HCN, as well as the Met−Met dimer. Note that for the smaller basis set, the lower pure DFT interaction energies mentioned above for the induction bonded dimers, combined with the smaller dispersion energies values, result in an error compensation for the total BH-B3LYP-D/6-311++G(2df,p) interaction energy. The largest underestimation errors for the aug-cc-pVTZ basis set are observed for the amino acid side chain dimers involving the aliphatic amino acid leucine (Leu), with the largest error found for the Leu−Leu dimer. A possible explanation is the underestimation of the multipole polarizabilities of this amino acid, requiring a larger basis set.

The underestimation of the dispersion energy values calculated with the smaller basis set are directly related to the smaller polarizability values, as can be seen from the molecular isotropic dipole, quadrupole, and octupole polarizability values summarized in Table 2. Herein, the isotropic polarizabilities are defined as $\alpha_{iso} = 1/3\Sigma_i\alpha_{ii}$, $\mathbf{C}_{iso} = 1/9\Sigma_{ij}\mathbf{C}_{ijij}$, and $\mathbf{R}_{iso} = 1/27\Sigma_{ijk}\mathbf{R}_{ijkijk}$. One can see that the polarizability values are significantly smaller when calculated using the smaller 6-311++G(2df,p) basis set. The underestimation of values is in particular large for the octupole polarizabilities, where on average only 70% of the value is recuperated using the smaller basis set. It is also noteworthy that the underestimation of the values is larger for the smallest systems, i.e., the rare gas atoms. However, the correlation between the polarizability values obtained using the two basis sets is very high, resulting also in a high correlation between the dispersion energies obtained using the two basis. This indicates that an extrapolation procedure to a complete basis set limit should be possible for the polarizabilities, allowing to obtain more accurate dispersion energies. Alternatively, one can adapt the approach developed by Chong,[51] who constructed specific basis set extensions, with a limited number of additional functions, for the specific goal of obtaining dynamic polarizabilities similar to values obtained at the complete basis set limits. Such procedures are not only desirable for reduction of the computational effort, as calculation of polarizabilities using large basis sets is expensive. They are also desirable in order to avoid numerical stability problems of the finite field procedure used to obtain the perturbed density matrices of the monomers: in larger systems, the convergence of the SCF procedure in the presence of an octupole field can become problematic, as was observed for several dimers from the S22[45] and SCAI[46] benchmark sets, excluded from the current study for this reason.

The contributions of the five calculated terms in the dispersion energy series for each examined dimer, calculated at the BH-B3LYP-D/aug-cc-pVTZ level, are summarized in Table 3 and graphically illustrated in Figure 3. In Figure 3, the dimers are ordered as in Table 3 from the lowest to the highest dispersion energies in absolute value and the contribution of each term in the series is displayed in percentage. For the smallest dimers, the $E_{disp}(R^{-6})$ term accounts for the largest part of the dispersion energy but the contribution of the $E_{disp}(R^{-8})$ and $E_{disp}(R^{-10})$ terms increases with the size of the system and the dispersion energy. In the largest systems, the $E_{disp}(R^{-10})$ contributes as much as 20%, and further study is necessary in order to determine whether this increasing trend will converge. It is possible that a damping function accounting for the effect of exchange on the dispersion energy will reduce the relative contribution of the higher terms. The size of the uneven terms $E_{disp}(R^{-7})$ and $E_{disp}(R^{-9})$ fluctuates as it is directly connected to the degree of the anisotropic character in the system: their values are equal to zero for the spherically symmetric rare gas dimers and become increasingly more important in the larger and more anisotropic

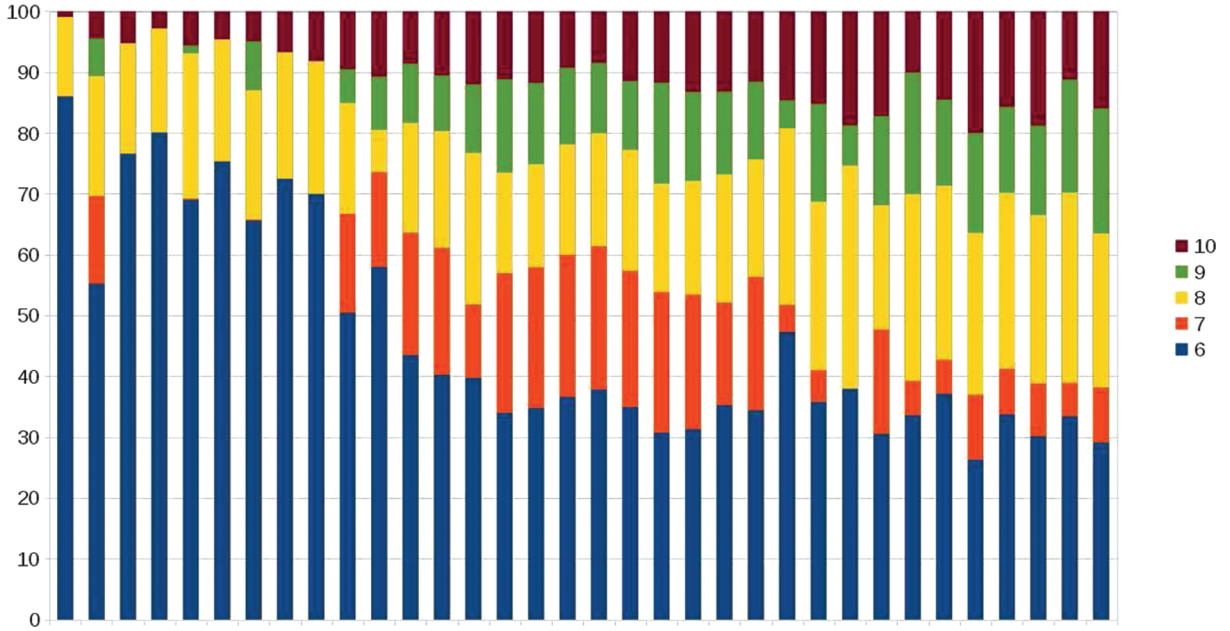

Figure 3. The contribution in percentage of the different terms $E(R^{-n})$, $n = 6-10$, to the total dispersion energy of the examined dimers. The dimers are ordered, as in Table 3, from the smallest to the largest dispersion energy in absolute value.

Table 2. The Isotropic Dipole (α_{iso}), Quadrupole (C_{iso}) and Octupole (R_{iso}) Polarizabilities (in au), Calculated Using the B3LYP Functional and the aug-cc-pVTZ and 6-311++G(2df,p) Basis Sets[a]

	α_{iso}		C_{iso}		R_{iso}	
monomer	aug-cc-pVZZ	6-311++G(2df,p)	aug-cc-pVTZ	6-311++G(2df,p)	aug-cc-pVTZ	6-311++G(2df,p)
He	1.47	0.64	0.99	0.43	2.05	0.13
Ne	2.52	1.49	2.91	2.24	6.08	3.04
Ar	11.14	7.89	19.78	15.36	97.27	48.36
N_2	12.06	11.03	40.91	30.30	145.03	103.21
FCl	18.36	14.73	62.62	45.66	319.62	209.96
CH_4	17.06	15.33	60.18	43.70	372.62	267.63
C_2H_2	23.87	20.92	96.18	76.10	585.48	392.15
C_2H_4	28.28	26.28	138.99	113.03	1198.09	955.59
C_6H_6	69.53	67.38	658.90	609.72	7801.84	7066.37
$C_4H_4N_2$	59.36	57.72	513.48	472.50	5407.16	4868.19
SiH_4	32.18	28.65	176.04	130.82	1502.42	1014.32
OCS	34.57	30.83	178.81	142.83	1758.48	1429.94
H_2O	9.77	7.60	21.76	14.70	115.41	63.27
NH_3	14.48	12.01	41.71	29.48	274.68	181.17
C_8H_7N	103.16	100.64	1392.48	1313.57	29082.92	27025.50
HCN	17.50	15.81	61.75	49.92	434.64	337.03
Ala	29.28	27.45	180.85	152.31	1644.78	1322.04
Leu	66.40	64.62	862.66	807.21	13697.17	12572.92
Ile	65.90	64.13	850.70	796.26	12723.07	11642.16
Gly	17.17	15.32	60.75	44.16	381.30	275.95
Thr	46.64	44.88	439.33	398.42	4932.69	4326.18
Met	76.11	73.39	1046.07	965.80	19445.89	17430.75
Val	53.96	52.34	554.09	510.07	7075.44	6355.31
R	0.9995		0.9997		0.9998	

[a] R is the correlation coefficient between the values calculated using the two basis sets.

system where they together contribute up to 40% of the dispersion energy. The relative effect of the anisotropy is expected to remain also after the inclusion of the damping function. For the most systems, one also observes a converging trend in the dispersion energy series

dx.doi.org/10.1021/ct200718y | J. Chem. Theory Comput. 2012, 8, 125–134

Table 3. Contribution of the Separate Terms $E_{disp}(R^{-n})$, $n = 6-10$, to the Dispersion Energy, Calculated Using the BH-B3LYP-D/aug-cc-pVTZ Method (All Values in kcal/mol)

dimer	$E_{disp}(R^{-6})$	$E_{disp}(R^{-7})$	$E_{disp}(R^{-8})$	$E_{disp}(R^{-9})$	$E_{disp}(R^{-10})$	E_{disp}
He_2	−0.03	0.00	0.00	0.00	0.00	−0.04
He−Ne	−0.05	0.00	−0.01	0.00	0.00	−0.06
He−Ar	−0.07	0.00	−0.02	0.00	0.00	−0.09
Ne_2	−0.07	0.00	−0.02	0.00	0.00	−0.10
He−N_2 L-shape	−0.06	−0.01	−0.02	−0.01	0.00	−0.10
He−N_2 T-shaped	−0.09	0.00	−0.03	0.00	−0.01	−0.13
He−FCl	−0.09	0.00	−0.03	−0.01	−0.01	−0.14
Ne−Ar	−0.12	0.00	−0.03	0.00	−0.01	−0.16
FCl−He	−0.19	−0.05	−0.02	−0.03	−0.04	−0.34
Ne−CH_4	−0.18	−0.06	−0.07	−0.02	−0.03	−0.36
Ar_2	−0.26	0.00	−0.08	0.00	−0.03	−0.37
Leu−Gly	−0.32	−0.15	−0.13	−0.07	−0.06	−0.74
CH_4−CH_4	−0.35	−0.18	−0.17	−0.08	−0.09	−0.88
CH_4−C_2H_4	−0.38	−0.11	−0.24	−0.11	−0.11	−0.95
Leu−Thr	−0.42	−0.28	−0.21	−0.16	−0.14	−1.22
Val−Leu	−0.42	−0.28	−0.21	−0.19	−0.14	−1.24
C_2H_4−C_2H_2	−0.49	0.01	−0.48	−0.08	−0.24	−1.28
Ile−Leu	−0.48	−0.31	−0.24	−0.16	−0.12	−1.31
Leu−Leu	−0.51	−0.31	−0.25	−0.15	−0.11	−1.34
SiH_4−CH_4	−0.55	−0.35	−0.31	−0.18	−0.18	−1.58
Ile−Ile	−0.59	−0.44	−0.34	−0.32	−0.23	−1.92
Ala−Leu	−0.64	−0.45	−0.38	−0.30	−0.27	−2.05
Val−Val	−0.76	−0.36	−0.46	−0.29	−0.29	−2.16
$(C_2H_4)_2$	−0.75	−0.48	−0.42	−0.28	−0.25	−2.18
$(OCS)_2$	−1.17	−0.11	−0.72	−0.11	−0.36	−2.47
C_6H_6−CH_4	−0.89	−0.13	−0.69	−0.40	−0.38	−2.50
C_6H_6−NH_3	−0.95	−0.14	−0.73	−0.36	−0.37	−2.56
C_6H_6−H_2O	−0.96	−0.21	−0.82	−0.40	−0.45	−2.84
C_6H_6−HCN	−0.96	−0.16	−0.90	−0.53	−0.32	−2.87
$(C_6H_6)_2$ T-shaped	−1.21	−0.20	−1.11	−0.72	−0.36	−3.62
Met−Met	−1.24	−0.70	−0.83	−0.59	−0.70	−4.05
C_6H_6−C_8H_7N T-shaped	−1.45	−0.45	−1.27	−1.03	−0.80	−5.00
$(C_4H_4N_2)_2$	−2.10	−0.60	−1.93	−1.02	−1.31	−6.96
$(C_6H_6)_2$ parallel-displaced	−1.91	−0.78	−1.94	−1.20	−1.46	−7.29

at the van der Waals minimum, where $E_{disp}(R^{-6}) > E_{disp}(R^{-8}) > E_{disp}(R^{-10})$, indicating that the effect of exchange at these geometries is not dominant, with the exception of the $(C_6H_6)_2$ parallel-displaced and the C_2H_4−C_2H_2 dimers, as discussed above.

5. SUMMARY AND CONCLUSIONS

In conclusion, we have presented a new combined Buckingham−Hirshfeld model (BH-DFT-D) for the evaluation of dispersion energies at the DFT level. By introducing a pairwise additive Coulomb interaction operator at the beginning of the derivation and defining distributed static multipole polarizability tensors within the framework of the iterative Hirshfeld method,[33,34,36] a four-centered dispersion energy correction to the interaction energy is obtained. The model constitues, as such, an improvement on our previous work,[17,18,23] due to the elimination of the pairwise additivity assumption of the dispersion energy. Furthermore, by making use of atomic polarizability tensors obtained from the *ab initio* molecular polarizabilities of the monomers, the full anisotropic character of the dispersion interaction is preserved. The model in its present form is suitable for application on dimers with no or only negligible overlap between the densities of the monomers.

The model is tested for a collection of 34 van der Waals dimers at equilibrium geometries by comparing the BH-B3LYP-D interaction energies with high level data obtained at the CCSD(T)/CBS level. The results obtained using the aug-cc-pVTZ basis set are found to be in good agreement with high level data, with a mean absolute error of 0.20 kcal/mol. The results obtained using the 6-311++G(2df,p) in general underestimate the dispersion energy due to the underestimation of the multipole polarizabilities of the monomers, in particular the octupole polarizability. The results can be further improved by extrapolation of the polarizability values to the complete basis set limit. The effect of the anisotropic terms $E(R^{-7})$ and $E(R^{-9})$ is found to be of increasing importance in the larger systems, contributing as much as 40% to the total dispersion energy value. The necessity of a damping function to compensate for the repulsive contribution of electron exchange to the dispersion energy is observed for several dimers, where the multipole expansion of the Coulomb

operator becomes divergent already at the van der Waals minimum. The next development steps of the model will consist of designing a damping function for shorter distances, derivation, or analytical gradients and generalization of the model to intramolecular dispersion interactions in macromolecules.

AUTHOR INFORMATION

Corresponding Author

*E-mail: alisa.krishtal@ua.ac.be.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This work was carried out using the Turing HPC infrastructure at the CalcUA core facility of the University of Antwerp, a division of the Flemish Supercomputer Center VSC, funded by the Hercules Foundation, the Flemish Government (department EWI), and the Universiteit Antwerpen. A.K. is grateful to the Research Foundation—Flanders (FWO) for a postdoctoral position and financial support.

REFERENCES

(1) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175.
(2) Pérez-Jordá, J. M.; Becke, A. D. *Chem. Phys. Lett.* **1995**, *233*, 134.
(3) Lotrich, V.; Bartlett, R. J. *J. Chem. Phys.* **2011**, *134*, 184108.
(4) Eshuis, H.; Furche, F. *J. Phys. Chem. Lett.* **2011**, *2*, 983–989.
(5) Dobson, J. F.; Dinte, B. P. *Phys. Rev. Lett.* **1996**, *76*, 1780–1783.
(6) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.
(7) Vydrov, O. A.; Voorhis, T. V. *Phys. Rev. Lett.* **2009**, *103*, 063004.
(8) Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010**, *82*, 092202.
(9) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
(10) Xu, X.; Goddard, W. A., III. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2673–2677.
(11) Zhang, Y.; Vela, A.; Salahub, D. R. *Theor. Chem. Acc.* **2007**, *118*, 693.
(12) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
(13) Sato, T.; Nakai, H. *J. Chem. Phys.* **2009**, *131*, 224104.
(14) Sato, T.; Nakai, H. *J. Chem. Phys.* **2010**, *133*, 194101.
(15) Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.
(16) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
(17) Krishtal, A.; Vannomeslaeghe, K.; Geldof, D.; Van Alsenoy, C.; Geerlings, P. *Phys. Rev. A* **2011**, *83*, 024501.
(18) Krishtal, A.; Vanommeslaeghe, K.; Olasz, A.; Veszprémi, T.; Van Alsenoy, C.; Geerlings, P. *J. Chem. Phys.* **2009**, *130*, 174101.
(19) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2007**, *127*, 154108.
(20) Kannemann, F. O.; Becke, A. D. *J. Chem. Theory Comput.* **2010**, *6*, 1081–1088.
(21) Steinmann, S. N.; Corminboeuf, C. *J. Chem. Theory Comput.* **2010**, *6*, 1990–2001.
(22) Steinmann, S. N.; Corminboeuf, C. *J. Chem. Phys.* **2011**, *134*, 044117.
(23) Olasz, A.; Vanommeslaeghe, K.; Krishtal, A.; Veszprémi, T.; Van Alsenoy, C.; Geerlings, P. *J. Chem. Phys.* **2007**, *127*, 224105.
(24) Jeziorski, B.; Moszynski, R.; Szalewic, K. *Chem. Rev.* **1994**, *94*, 1887–1930.
(25) Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2003**, *367*, 778–784.
(26) Heßelmann, A.; Jansen, G.; Schtz, M. *J. Chem. Phys.* **2005**, *122*, 014103.
(27) Misquitta, A. J.; Szalewicz, K. *J. Chem. Phys.* **2005**, *122*, 214109.
(28) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2005**, *123*, 214103.
(29) Rajchel, Ł.; Żuchowski, P. S.; Szczesniak, M. M.; Chałasinski, G. *Phys. Rev. Lett.* **2010**, *104*, 163001.
(30) Buckingham, A. D. *Adv. Chem. Phys.* **1967**, *12*, 107.
(31) London, F. *Trans. Faraday Soc.* **1937**, *33*, 8.
(32) Stone, A. J.; Tong, C. S. *Chem. Phys.* **1989**, *137*, 121–135.
(33) Hirshfeld, F. L. *Theor. Chim. Acta* **1977**, *44*, 129–139.
(34) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbo-Dorca, R. *J. Chem. Phys.* **2007**, *126*, 144111.
(35) Nalewajski, R. F.; Broniatowska, E. *Int. J. Quantum Chem.* **2005**, *101*, 3957.
(36) Krishtal, A.; Senet, P.; Yang, M.; Van Alsenoy, C. *J. Chem. Phys.* **2006**, *125*, 034312.
(37) Oláh, J.; Blockhuys, F.; Veszprémi, T.; Van Alsenoy, C. *Eur. J. Inorg. Chem.* **2006**, 69–77.
(38) Krishtal, A.; Vyboishchikov, S.; Van Alsenoy, C. *J. Chem. Theory Comput.* **2011**, *7*, 2049–2058.
(39) Vanfleteren, D.; Van Neck, D.; Bultinck, P.; Ayers, P.; Waroquier, M. *J. Chem. Phys.* **2010**, *132*, 164111.
(40) Vanfleteren, D.; Van Neck, D.; Bultinck, P.; Ayers, P.; Waroquier, M. *J. Chem. Phys.* **2011**, *133*, 231103.
(41) Hättig, C.; Jansen, G. H.; Hess, B. A.; Ángyán, J. G. *Mol. Phys.* **1997**, *91*, 145–160.
(42) Geldof, D.; Krishtal, A.; Geerlings, P.; Van Alsenoy, C. *J. Chem. Phys. A* **2011**, *115*, 13096–13103.
(43) Le Sueur, C. R.; Stone, A. J. *Mol. Phys.* **1993**, *78*, 1267.
(44) Bader, R. F. W. *Chem. Rev.* **1991**, *91*, 893.
(45) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
(46) Berka, K.; Laskowski, R.; Riley, K. E.; Hobza, P.; Vondrasek, J. *J. Chem. Theory Comput.* **2009**, *5*, 982–992.
(47) 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, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
(48) Rousseau, B.; Peeters, A.; Van Alsenoy, C. *Chem. Phys. Lett.* **2000**, *324*, 189.
(49) Van Alsenoy, C.; Peeters, A. *THEOCHEM* **1993**, *105*, 19.
(50) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(51) Chong, D. P. *Mol. Phys.* **2005**, *103*, 749.