

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/5896744>

Parametrization of an Empirical Correction Term to Density Functional Theory for an Accurate Description of π -Stacking Interactions in Nucleic Acids

ARTICLE in THE JOURNAL OF PHYSICAL CHEMISTRY B · DECEMBER 2007

Impact Factor: 3.3 · DOI: 10.1021/jp072317s · Source: PubMed

CITATIONS

37

READS

14

2 AUTHORS, INCLUDING:



Jean-Marie Ducéré

University of Toulouse

29 PUBLICATIONS 284 CITATIONS

SEE PROFILE

Parametrization of an Empirical Correction Term to Density Functional Theory for an Accurate Description of π -Stacking Interactions in Nucleic Acids

Jean-Marie Duc  re and Luigi Cavallo*

Dipartimento di Chimica, Universit   di Salerno, Via ponte don Melillo, Fisciano (SA), I-84084, Italy

Received: March 23, 2007; In Final Form: August 20, 2007

We present an accurate parametrization of density functional theory augmented with an empirical correction term to describe properly π -stacking interactions in nucleic acids. The approach is based on the popular Perdew–Burke–Ernzerhof (PBE), Becke–Perdew (BP), and hybrid Becke–Lee–Yang–Parr (B3LYP) density functionals augmented by a classical London C_6R^{-6} dispersion term. The novelty of our implementation lies in the accurate tuning of the empirical parameters, included in the $\gamma_d(R)$ damping function, to reproduce high-level post Hartree–Fock calculations. In particular, we present sets of parameters and the needed code to correct the PBE, BP, and B3LYP results from the Turbomole and ADF packages in connection with basis sets of double and triple ζ quality. The developed approach is validated by comparison with the JSCH-2005 benchmark and with best quality stacking energies reported in the literature for the stacking of H-bonded nucleic acids base pairs.

Introduction

Van der Waals (or dispersion) interactions play a major role in many chemical systems. They are important for the packing of molecular crystals, for host–guest interactions and, more generally, for all systems involving neutral, non-, or weakly-polar species. In particular, they are fundamental to the stability of systems with stacked aromatic rings and thus are of paramount relevance for the stability of DNA and RNA. Therefore, any computational method used for dealing with such systems should properly account for dispersion interaction. However, while some molecular mechanics force-fields can give reasonable results, not all quantum mechanics methods perform well.

On one hand, very high level correlated wavefunction methods, like extrapolated coupled-clusters calculations with inclusion of single, double, and perturbatively connected triple excitations [CCSD(T)] give good results. The main drawback of such high level methods is their prohibitive computational cost that forbids their use for systems larger than some atoms. The computationally much more affordable second order M  ller–Plesset theory method (MP2) can be used as an alternative, at the expense of the accuracy. Dispersion interactions relevant to aromatic stacking are rather systematically overestimated by this method, even if in some cases lucky error cancellations can occur by choosing an adequate basis set, like the 6-31G*(0.25) basis set proposed by Hobza et al. that performs well for aromatic stacking.¹ Still, MP2 calculations remain computationally too demanding for calculating geometries of systems larger some tens of atoms, even when using the Resolution of the Identity approximation to speed up the calculations.^{2,3}

On the other hand, Hartree–Fock (HF) and density functional theory (DFT) methods are much less computationally demanding and allow the treatment of systems of hundreds of atoms. However, the HF method cannot cover dispersion interaction

since this kind of interaction arises from electron correlation. In the case of DFT, electron correlation is treated in an approximate way depending on the functional. Yet, results are very often disappointing, except in some cases where errors fortuitously cancel. According to Parrinello et al., all standard density functionals, as they rely on the local electron density, its gradient, and the local kinetic energy density, cannot properly deal with dispersion interaction.⁴ Different ways to improve density functionals to properly treat dispersion interaction have been investigated,^{5,6} but according to Grimme, it is likely that this approach will end in methods at least as complex (and demanding) as MP2.⁷

To date, the most practical and successful approach to cover for dispersion interactions has been adding an empirical correction term to HF^{8,9} and DFT.^{4,7,10–14} The correction energy used, E_{disp} , is of the form $E_{\text{disp}} = \gamma_d(R)C_6R^{-6}$, where $\gamma_d(R)$ is an appropriate damping function needed to switch off the E_{disp} term at short distances (vide infra), and C_6R^{-6} is a classical London-type term. Such correction allows one to correct the long range behavior of the functionals and has been applied to areas as diverse as small van der Waals complexes,^{4,7,11,15,16} biomolecules,^{12–14,17–19} organic chemistry,²⁰ or material sciences,^{21–23} with encouraging results.

However, in most of these works the parameters of the damping function are chosen quite arbitrarily, such as by matching the pioneering Ahlrichs’s function^{8,9} or by fixing the steepness of the damping function, but for few valuable exceptions,^{4,7,11,13,15,16} the effect of the basis set or the functional was not investigated. This implies that in most cases the quantum mechanics term and the empirical correction term were added without a proper tuning of all of the parameters included in the damping function $\gamma_d(R)$, which is of course needed since the quantum mechanics term is dependent on the specific implementation used (HF, DFT, and different functionals) and on the basis set used.

In this article, we propose an empirical method to fit all of the parameters of the damping function $\gamma_d(R)$ in E_{disp} to reproduce high-level post-HF MP2 or CCSD(T) calculations. This way

* Corresponding author. E-mail lcavallo@unisa.it.

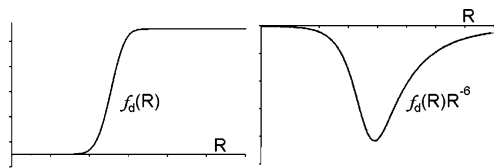


Figure 1. Shape of the damping function $f_d(R)$ (left) and of damped R^{-6} (right).

we parametrized a set of damping functions to correct the weak performances of three specific and largely used functionals, such as the Perdew–Burke–Ernzerhof (PBE),²⁴ the Becke–Perdew (BP),^{25–27} and the hybrid Becke–Lee–Yang–Parr (B3LYP)^{25,28,29} functionals. In particular, we worked on the Turbomole³⁰ implementation of the three functionals in connection with the SVP and TZVP basis sets,^{31,32} as well as on the ADF^{33–35} implementation of the PBE and BP functionals in connection with the DZP and TZP basis sets.^{33–35} We tested our approach to reproduce the interaction energies of the JSCH-2005 benchmark.³⁶ This is a severe test since the JSCH-2005 benchmark is based on a database of CCSD(T) complete basis set limit intermolecular interaction energies of 165 noncovalent complexes (128 nucleic acid base pairs, 19 amino acid pair, and 18 small complexes; Supporting Information). We further tested our approach to reproduce CCSD(T) complete basis set limit stacking energy between some nucleic base pairs reported by Sponer et al.,³⁷ since these systems offer the opportunity to test the approach with larger systems (four interacting fragments) than those included in the JSCH-2005 database (two interacting fragments). The program needed to integrate the Turbomole package is available upon request, while the official ADF release already includes our modifications.

We focused specifically on nucleic acids since an accurate simulation of RNA and DNA structures is a field of indubitable relevance. In this respect, correct reproduction of structure and energetics of nucleic acids requires the following effects to be correctly reproduced: (1) H-bonding interactions, (2) solvent effects, and (3) stacking interactions. While quantum mechanics approaches are accurate enough in evaluating the first two terms,^{38,39} the failure of standard approaches in reproducing stacking interactions severely limits the applicability of quantum mechanics simulations to a field of large chemical and biological interest. This is the reason for focusing on an accurate parametrization for simulating nucleic acids specifically.

Models and Methods

Empirical Correction. The analytical form assumed for the dispersion energy correction is the now established form of eq 1:

$$E_{\text{disp}} = - \sum_i \sum_j f_d(R_{ij}) \times \frac{C_6^{ij}}{R_{ij}^6} \quad (1)$$

where the E_{disp} is expressed as the sum of diatomic ij contributions. The $f_d(R_{ij})$ function (Figure 1) is continuous and has continuous derivatives with respect to R_{ij} . Its role is to damp the R^{-6} term at short distances, where the calculation method, here, the density functional, is thought to perform well.

For obtaining the C_6 parameters, we followed the method proposed by Elstner et al.¹⁰ They were derived using the Slater–Kirkwood (SK) approximation of eq 2.^{40,41}

$$C_6^{ij} = \frac{3}{2} \frac{\alpha_i \alpha_j}{\sqrt{\alpha_i/N_i} + \sqrt{\alpha_j/N_j}} \quad (2)$$

TABLE 1: Miller’s Atomic C_6 and Polarizability α , from Ref 42, and Bondi Radius from Ref 44^a

	h	cte	ctr	cbr	npi2	ntr2	otr4
C_6	2.83	22.7	32.6	54.17	25.13	23.08	10.02
α	2.61	7.16	9.12	12.80	7.36	6.95	3.84
R_0	2.27	3.21	3.21	3.21	2.93	2.93	2.87

^a All values are in atomic units. Atom types from ref 42: h = all H atoms; cte = sp^3 C atoms; ctr = sp^2 C atoms; cbr = ring-junction sp^2 C atoms; npi2 = trigonal sp^2 N atoms; ntr2 = sp^2 N atoms; otr4 = sp^2 O atoms.

that is more commonly written as eq 3:

$$C_6^{ij} = \frac{2C_6^{ii}C_6^{jj}\alpha_i\alpha_j}{C_6^{ii}\alpha_j^2 + C_6^{jj}\alpha_i^2} \quad (3)$$

where α_i and N_i are the polarizability and the effective number of electrons of atom or molecule i . By taking $i = j$ in eq 2 gives a route to the like-atoms atomic parameters, C_6^{ii} , of eq 4:

$$C_6^{ii} = \frac{3}{4} \sqrt{\alpha_i^3 N_i} \quad (4)$$

In the present work, we used the atomic polarizabilities proposed by Miller, fitted to reproduce the molecular ones; see Table 1.⁴² These polarizabilities depend on the element and its hybridization state. For the effective number of electrons, we used the formula proposed by Halgren:⁴³ $N_i = 0.8$ for H atom, and $N_i = 1.17 + 0.33n_i$ for C–Ne atoms, where n_i is the number of valence electrons of atom i .

Having the atomic polarizabilities and effective numbers of electrons, we can use eq 2 or, alternatively, eqs 4 and 3 to calculate all of the C_6 parameters required in eq 1.

The other part of eq 1 is the damping function $f_d(R)$. We opted for the simple sigmoid expression reported in eq 5:

$$f_d(R_{ij}) = \frac{1}{1 + \exp(-a(R_{ij} - R_0))} \quad (5)$$

where R_0 is a parameter related to the size of the atoms paired and a is the steepness of the function. This function is very similar to the Fermi function used by Yang.¹¹ The difference is that in our case the steepness is independent of the size of the atoms (R_0). We took R_0 as the sum the Bondi’s radii⁴⁴ of atoms i and j , but this is not the only possible choice. For example, Grimme derived his radii from atomic HF calculations. To improve flexibility in our parametrization scheme and to allow the use of any other kind of (consistent) radii, we scaled R_0 by an additional parameter b , to be fitted, as shown in eq 6:

$$f_d(R_{ij}) = \frac{1}{1 + \exp(-a(R_{ij} - bR_0))} \quad (6)$$

Finally, since different combinations of density functionals and basis sets perform differently when it comes to dispersion interaction, we followed Grimme’s approach and introduced the additional parameter c , to be fitted, that depends on the specific functional and basis set used.⁷ This leads to the final expression of the damping function; see eq 7:

$$f_d(R_{ij}) = \frac{c}{1 + \exp(-a(R_{ij} - bR_0))} \quad (7)$$

By replacing $f_d(R_{ij})$ from eq 7 into eq 1, we obtain the final expression of E_{disp} shown in eq 8:

$$E_{\text{disp}} = - \sum_i \sum_j \frac{c}{1 + \exp(-a(R_{ij} - bR_0))} \times \frac{C_6^{ij}}{R_{ij}^6} \quad (8)$$

Gradients with respect to nuclear displacements are obtained easily by applying the chain rule to eq 8; see eq 9:

$$\frac{\partial E_{\text{disp}}}{\partial x_i} = \frac{\partial E_{\text{disp}}}{\partial R_{ij}} \frac{\partial R_{ij}}{\partial x_i} = \sum_j \frac{x_i - x_j}{R_{ij}} \times \mathcal{f}_d(R_{ij}) \times \frac{C_6^{ij}}{R_{ij}^6} \times \left(a - \frac{a}{c} \times \mathcal{f}_d(R_{ij}) - \frac{6}{R_{ij}} \right) \quad (9)$$

where $\mathcal{f}_d(R_{ij})$ is the damping function of eq 7.

We note that the difference between our expression of E_{disp} and the expression proposed by Grimme is in the $\exp(-a(R_{ij} - bR_0))$ term, which replaces the $\exp(-a(R_{ij}/R_0 - 1))$ term used by Grimme.⁷

As the correction used is intrinsically empirical, there is no really rigorous way to determine the a , b , and c parameters. The method we adopted was to fit these parameters to reproduce correlated post-Hartree–Fock data as these methods account for the dispersion interaction. Finally, it is worth noting that what we call E_{disp} is not the true dispersion energy. Indeed, E_{disp} is rather a correction to account for missing or badly described dispersion energy in the density functional. We found it convenient not to correct the ΔE_{DFT} for the BSSE but to include it into E_{disp} , which is the procedure commonly adopted in this kind of calculation. While leading to a somewhat unclear definition of what E_{disp} really is, this approach provides “BSSE-free” geometries. This can present a great advantage, since it is well-known that BSSE can deteriorate the geometries of weakly bound complexes significantly (see for example BSSE-corrected and -uncorrected results in ref 45). However, the price to pay for this convenience is the need to reparameterize the damping function for each basis set. Of course, this approach also eliminates the need of a posteriori BSSE correction.

Computational Details. The MP2 calculations were performed using the Turbomole 5.8 package³⁰ using the aug-cc-pVTZ basis set.^{46,47} The resolution of the identity was used for both the HF (RI-JK)⁴⁸ and the MP2^{2,3} steps. As there is no auxiliary basis set corresponding to aug-cc-pVTZ for RI-JK, we used the one corresponding to cc-pVTZ. Comparisons with non-RI HF calculations showed differences of less than 0.1 kcal/mol for the interaction energies. All of the MP2 interaction energies were corrected for the BSSE using the Boys and Bernardi counterpoise scheme (CP in the following).⁴⁹

The DFT calculations were performed with a version of ADF 2005.01^{33–35} modified by us and with the Turbomole 5.8 package, to which we add our dispersion correction with a stand-alone program to be run between the rdgrad and relax Turbomole programs. ADF calculations were run using the DZP and TZP basis sets keeping frozen the 1s core of C, N, and O atoms. We will refer to these calculations as ADF/DZP and ADF/TZP, respectively. Turbomole calculations were done using the SVP and TZVP basis sets^{31,32} and making use of the RI-J^{50,51} approximation. We will refer to these calculations as TM/SVP and TM/TZVP, respectively.

Fitting Procedure. A first training data set was constituted of the potential energy curves of adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U) stacked dimers. We investigated the interaction energy dependence with the twist angle and the vertical separation. The twist angle dependence

was examined fixing the distance between the base planes at 3.3    for A–A, T–T, and U–U and at 3.4    for C–C and G–G, and rotating about the axis passing through the center of mass of the two bases from 0   to 180   (Figure 2), similar to ref 11. A twist angle of 0   corresponds to two perfectly stacked bases. The vertical separation was studied for a twist angle of 180   by varying the distance between the bases from 2.8 to 3.8   . The geometries of the monomers were obtained by geometry optimizations using the PBE functional with the SVP basis set. The potential energy curves of the different dimers were built using the MP2 method and the aug-cc-pVTZ basis set (cf. computational details). The main advantage of this rather large training data set is that these systems are similar to the systems relevant in the field of biochemistry. The main drawback is that the MP2 method overestimates this kind of dispersion interaction. This training data set was fitted with the PBE functional only for the ADF/DZP, ADF/TZP, TM/SVP, and TM/TZVP combinations of programs and basis sets.

As an alternative training data set, we used the potential energy curves reported by Sherrill et al.⁵² for the benzene dimer. These curves describe the dependence of the interaction energy with the inter-monomer distance for the three isomers of the benzene dimer reported in Figure 3. The geometry of the monomer was kept frozen in Gauss et al. geometry,⁵³ which means $R_{\text{C–C}} = 1.3915$    and $R_{\text{H–H}} = 1.0800$   . These curves were calculated at the MP2 level using a modified aug-cc-pVQZ basis set. In addition, they include a CCSD(T) correction estimated using the smaller aug-cc-pVDZ basis set. This calculation method is not far from providing the most accurate data possible. The main advantage of this training data set is that dispersion interaction is described with good accuracy. The main drawback is that the training data set is limited and that a benzene dimer is not particularly representative of the biological systems of interest. This training data set was fitted using the three PBE, PB, and B3LYP functionals, and only the ADF/TZP and TM/TZVP combinations of programs and basis sets were considered. Since the benzene dimer interaction energy is small, the large BSSE due to the use of double ζ basis sets would lead to less reliable and meaningful results.

In all cases, the parameters optimization was carried out by minimizing the value of σ , defined in eq 10:

$$\sigma = \sum_{\text{training set}} (\Delta E_{\text{ab initio}} - \Delta E_{\text{DFT–D}})^2 = \sum_{\text{training set}} (\Delta E_{\text{ab initio}} - (\Delta E_{\text{DFT}} + \Delta E_{\text{Disp}}))^2 \quad (10)$$

where $\Delta E_{\text{ab initio}}$ is the interaction energy between the two nucleic acid bases at the MP2 level in the case of the first training data set or the CCSD(T) interaction energy between the two benzene molecules in the case of the second training data set. $\Delta E_{\text{DFT–D}}$ is the van der Waals corrected DFT interaction energy, while ΔE_{DFT} and ΔE_{Disp} are the DFT and dispersion contributions to the total $\Delta E_{\text{DFT–D}}$ interaction energy. The term training set refers either to the MP2 potential energy curves of the nucleic acid bases dimer or to the CCSD(T) potential energy curves of the benzene dimer. In addition, the damping function was constrained to go to zero (or at least to nearly negligible values) within the range of covalent bonds (1–1.5   ) where DFT has been shown to perform well. We will refer to these corrections as Disp^{MP2} when using the damping function optimized on the nucleic bases training data set and as $\text{Disp}^{\text{CCSD(T)}}$ when using the benzene training data set. We will refer to the corresponding dispersion energies as $E_{\text{Disp}}^{\text{MP2}}$ and $E_{\text{Disp}}^{\text{CCSD(T)}}$, respectively.

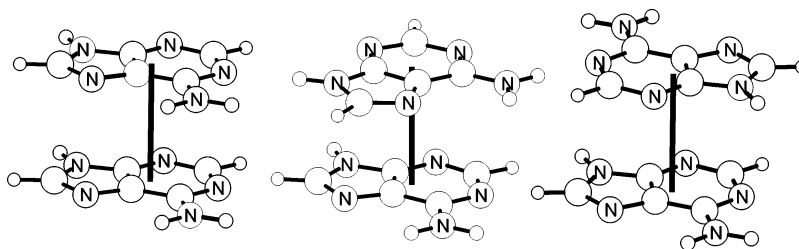


Figure 2. Adenine dimer for a twist angle of 0 (left), 60 (middle), and 180° (right).

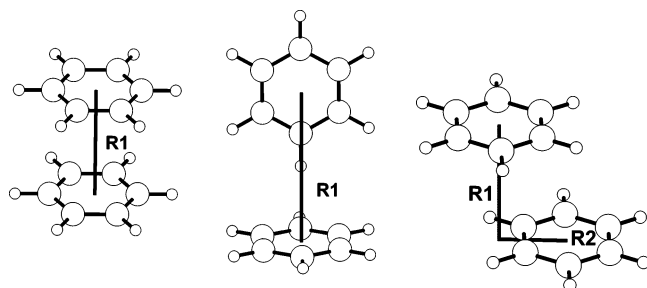


Figure 3. Isomers of the benzene dimer: sandwich (left), T-shaped (middle), parallel displaced (right). R1 and R2 are the parameters varied in the potential energy curves.

Results and Discussion

Fitting of the MP2 Nucleic Base Dimers Energy Curves.

As an exemplary case of the well-known poor performances of DFT (the PBE functional in this case) to capture dispersion interactions, the BSSE uncorrected PBE curves of the adenine dimer are reported in Figure 4. In the same figure, the CP corrected MP2 curves are also reported.

Visual inspection clearly indicates that at the MP2 level the adenine dimer is remarkably stable in the gas phase. Differently, both of the triple ζ basis set approaches (ADF/TZP and TM/TZVP) give very similar and completely repulsive curves. Because of the moderate BSSE error, the TM/SVP double ζ basis set results in substantially noninteracting nucleic acid bases, while the rather larger BSSE associated with the ADF/DZP approach partly mimics the dispersion interaction leading to an attractive curve. Very similar behavior was found for all of the other nucleic acid base dimers.

Minimization of σ of eq 10 leads to different sets for the a , b , and c parameters of eq 7, one for each of the DFT-based approaches (program and basis set) used here. After inclusion of the E_{disp} contribution, and independent of the DFT approach, the resulting $\Delta E_{\text{DFT-D}}$ curves are very close to the MP2 curves (see Figure 5), with a maximum error of about 1.5 kcal/mol.

The largest deviations occur at short inter-base separation, ≤ 2.8 Å, which is well below the typical range for base stacking.

Fitting of the CCSD(T) Benzene Dimer Energy Curves.

The BSSE uncorrected PBE curves of the benzene dimer, E_{DFT} , as well as the same curves corrected for the dispersion contribution, $E_{\text{DFT-D}}$, are reported in Figure 6. In the same figure, the CP corrected CCSD(T) curves are also reported. Also, in this case, the simple E_{DFT} curves fail to reproduce the post-HF calculations. Only for the T-shaped geometry can a moderately attractive behavior be found at the uncorrected PBE level. Fitting the parameters a , b , and c of eq 10 for the ADF/TZP and TM/TZVP approaches leads to $E_{\text{DFT-D}}$ curves. After inclusion of the E_{disp} correction, the difference between the CCSD(T) and the $E_{\text{DFT-D}}$ curves is less than 0.4 kcal/mol. Very similar conclusions are reached when the BP and B3LYP functionals are used. For the sake of simplicity these curves are not discussed in details.

The values of the optimized parameters for all of the functionals used are reported in Table 2. The a parameter is rather independent of the training data set and of the basis set, with the exception of the ADF/DZP basis set, but strongly depends on the functional. The b parameter, instead, is consistently around 1.0. On the other hand, the c parameter, which corresponds to the value of the damping function at large distances, is very sensitive to both the basis set and the training data set and, as already indicated by Grimme, is also very sensitive to the functional.⁷ With the exception of the c parameter connected to the TM/SVP approach, the fitted c parameters for the PBE functional are in the range 0.6–0.8, in good agreement with the parameter 0.7 proposed by Grimme.⁷ The main conclusion of this section is that parameters cannot be transferred between different functionals and/or basis sets without careful tests.

Validation of the Optimized Parameters. To validate the approach we developed, we tested it with the JSCH-2005 benchmark. Statistical deviations are reported in Table 3. For

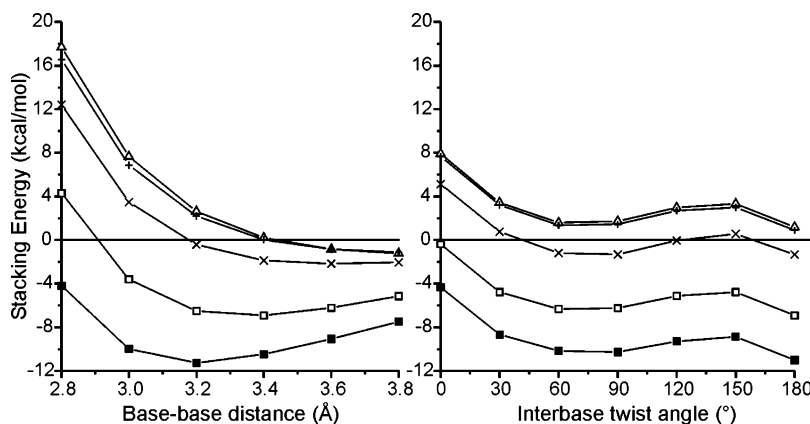


Figure 4. Uncorrected potential energy curves for the adenine dimer. Vertical separation (left), and twist angle (right) dependence. \square , ADF/DZP; Δ , ADF/TZP; \times , TM/SVP; $+$, TM/TZVP; \blacksquare , MP2. The DFT curves are obtained with the PBE functional.

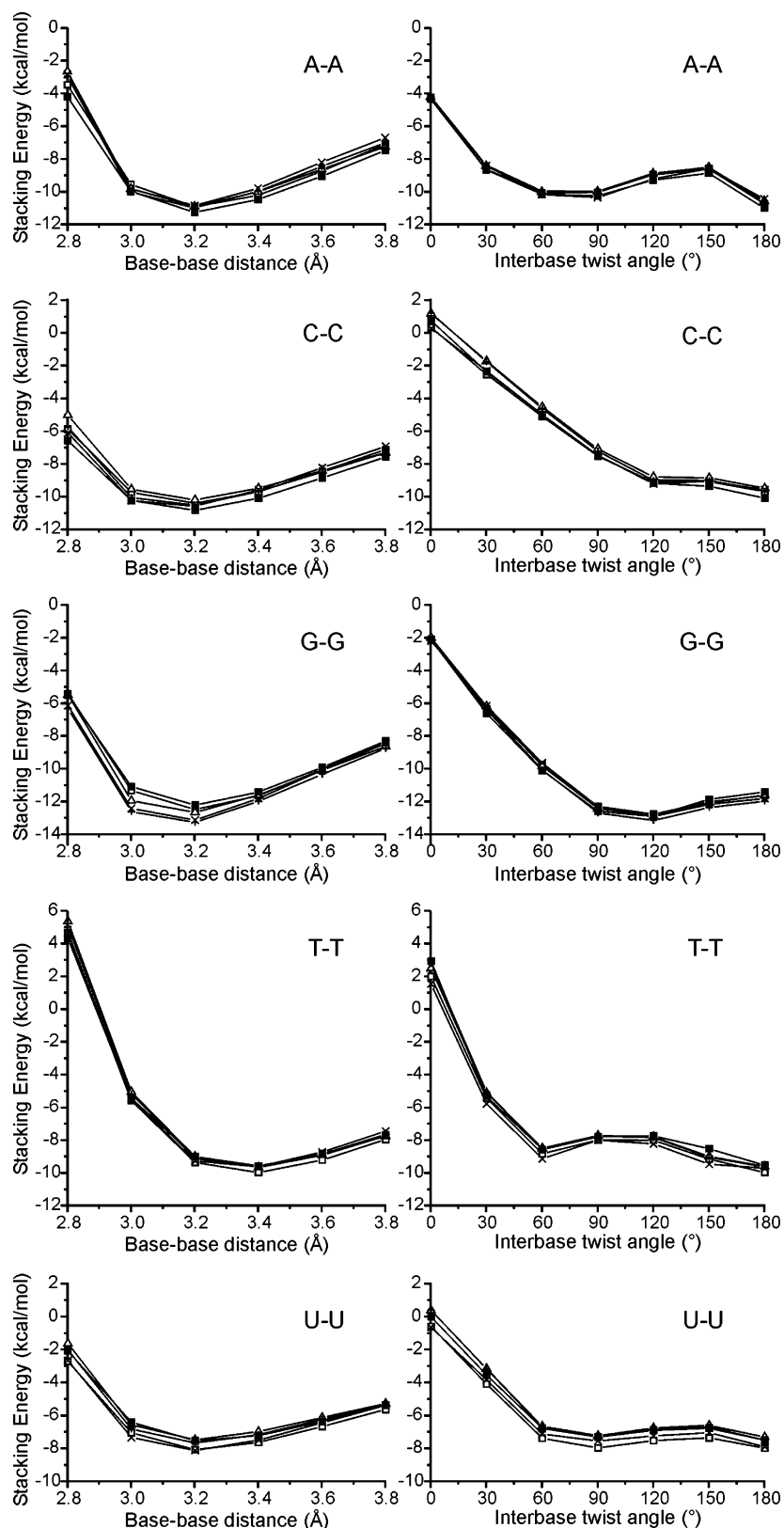


Figure 5. Vertical separation (left) and twist angle (right) dependence energy curves for nucleic bases including $E_{\text{Disp}}^{\text{MP2}}$ correction to DFT calculation.   , ADF/DZP + $E_{\text{Disp}}^{\text{MP2}}$;   , ADF/TZP + $E_{\text{Disp}}^{\text{MP2}}$;   , TM/SVP + $E_{\text{Disp}}^{\text{MP2}}$; +, TM/TZVP + $E_{\text{Disp}}^{\text{MP2}}$;   , MP2. The DFT curves are obtained with the PBE functional.

the sake of comparison, we also report the deviations without E_{Disp} correction. Further, we tested our approach with the larger TZV2P and TZ2P basis sets for the TURBOMOLE and ADF calculations, respectively. In this case, we applied the $E_{\text{Disp}}^{\text{CCSD(T)}}$ correction term obtained with the parameters developed for the

same programs, and we scaled the c parameter of eq 7 in order to minimize the MAD deviation from the reference JSCH-2005 benchmark.

As expected, pure DFT approaches remarkably deviate from the reference interaction energies of the JSCH-2005 benchmark

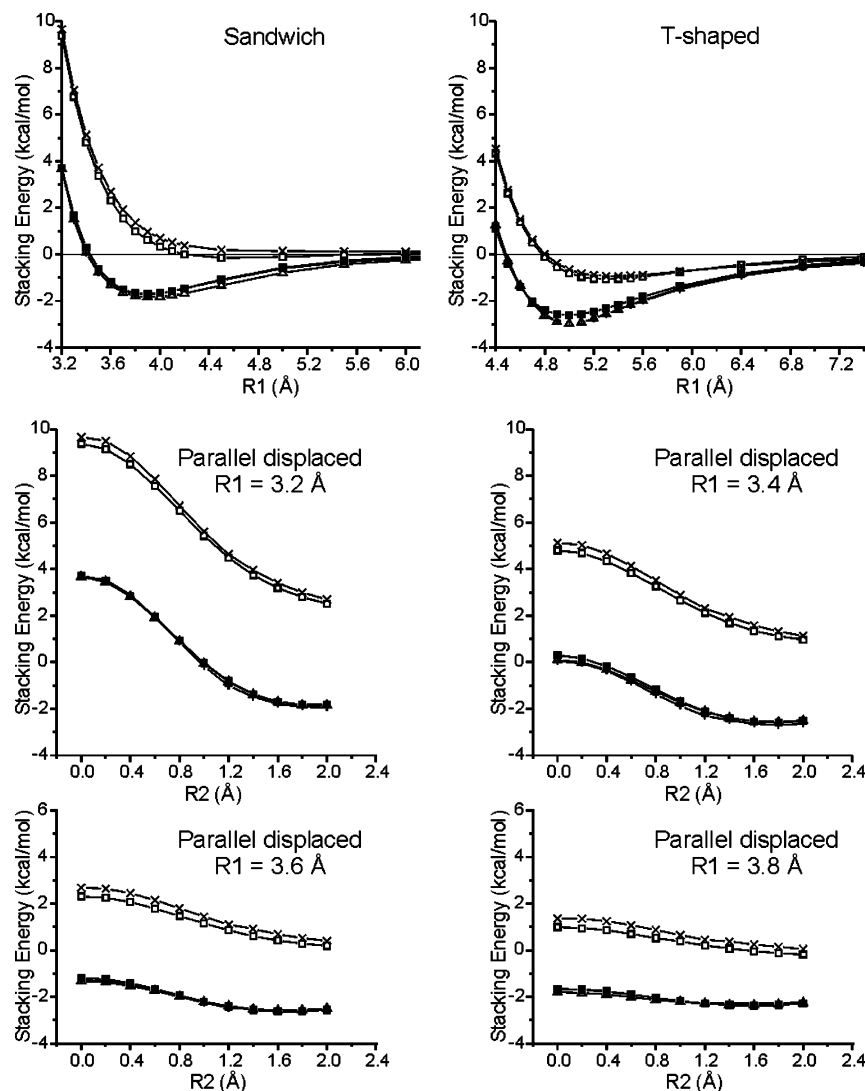


Figure 6. Potential energy curves of the benzene dimer isomers. \square , ADF/TZP; \triangle , ADF/TZP + $E_{\text{Disp}}^{\text{CCSD(T)}}$; \times , TM/TZVP; $+$, TM/TZVP + $E_{\text{Disp}}^{\text{CCSD(T)}}$; \blacksquare , MP2/CCSD(T). The DFT curves are obtained with the PBE functional.

TABLE 2: Optimized Values of the a , b , and c Parameters

program/basis set	functional	training set					
		nucleic base dimers (MP2 curves)			benzene dimer (CCSD(T) curves)		
		a (\AA^{-1})	B	c	a (\AA^{-1})	b	c
ADF/DZP	PBE	18.40	0.8342	0.2662			
ADF/TZP	PBE	10.82	0.8710	0.8333	8.906	0.9343	0.6496
ADF/TZP	BP				1.134	1.0501	1.3782
TM/SVP	PBE	10.05	0.8840	0.6364			
TM/TZVP	PBE	10.26	0.8806	0.8029	8.907	0.9506	0.7133
TM/TZVP	BP				1.089	1.0869	1.5433
TM/TZVP	B3LYP				1.355	0.9556	1.2185

(rmsd > 3 kcal/mol). The only exception is the PBE/ADF/DZP approach, that shows a rather small rmsd, 1.35 kcal/mol. However, this reasonable performance is the consequence of the high BSSE between stacked basis previously discussed. Inclusion of the dispersion correction remarkably reduces deviation from the JSCH-2005 benchmark, with a rmsd usually comprised between 1 and 2 kcal/mol. The only exception is the PBE/TM/SVP + Disp^{MP2} approach, with a rmsd of 3.69 kcal/mol. In both the TM/TZVP and the ADF/TZP approaches, the $E_{\text{Disp}}^{\text{CCSD(T)}}$ correction term overperforms the $E_{\text{Disp}}^{\text{MP2}}$ one, which validates an approach that was parametrized with a rather small training data set of high quality data. Overall, the performance of the PBE abd BP TM/TZVP + Disp^{CCSD(T)} and ADF/TZP +

Disp^{CCSD(T)} approaches is very similar. Using the larger TZV2P and TZ2P basis sets and scaling the c parameter (by 0.97 and 1.06 for the PBE/TM/TZV2P and PBE/ADF/TZ2P approaches, respectively) improves the rmsd slightly, in agreement with previous results by Grimme.¹⁴ A comparison between the three functionals indicates that the PBE functional performs slightly better than the BP and B3LYP functionals, which result in slightly higher rmsd.

Finally, we also report statistic data for the whole benchmark, where the dispersion interaction is not added to the H-bonded complexes, that is, the hydrogen-bonded base pairs 1–38, and complexes 140, 145–150, and 165 of the JSCH-2005 database. The data reported in Table 3 indicate that removing dispersion

TABLE 3: Deviations, in kcal/mol, from the Complete (143 Complexes + S22 Dataset) JSCH-2005 Benchmark Database^a

program/ basis set	functional	dispersion term fitted to ^b	E_{Disp} on H-bonded systems	MD	MAD	RMSD
TM/SVP	PBE			1.06	2.87	3.51
TM/TZVP	PBE			3.38	3.47	4.51
TM/TZV2P	PBE			3.63	3.70	4.73
TM/TZVP	BP			5.27	5.29	6.41
TM/TZVP	B3LYP			4.89	4.92	6.03
ADF/DZP	PBE			0.33	0.98	1.35
ADF/TZP	PBE			3.63	3.69	4.76
ADF/TZ2P	PBE			3.72	3.77	4.83
ADF/TZP	BP			5.54	5.56	6.71
TM/SVP	PBE	MP2	yes	-2.78	2.78	3.69
TM/TZVP	PBE	MP2	yes	-1.48	1.57	1.97
ADF/DZP	PBE	MP2	yes	-1.33	1.47	1.76
ADF/TZP	PBE	MP2	yes	-1.46	1.55	1.94
TM/TZVP	PBE	CCSD(T)	yes	-0.60	0.87	1.26
TM/TZV2P	PBE	CCSD(T)	yes	-0.35	0.74	1.07
TM/TZVP	BP	CCSD(T)	yes	-0.77	1.07	1.56
TM/TZVP	B3LYP	CCSD(T)	yes	-1.27	1.38	1.83
ADF/TZP	PBE	CCSD(T)	yes	-0.08	0.81	1.17
ADF/TZ2P	PBE	CCSD(T)	yes	0.01	0.79	1.15
ADF/TZP	BP	CCSD(T)	yes	-0.28	0.97	1.40
TM/SVP	PBE	MP2	no ^c	-1.96	1.97	2.39
TM/TZVP	PBE	MP2	no ^c	-0.43	1.18	1.45
ADF/DZP	PBE	MP2	no ^c	-0.95	1.14	1.41
ADF/TZP	PBE	MP2	no ^c	-0.34	1.22	1.46
TM/TZVP	PBE	CCSD(T)	no ^c	0.17	0.75	1.05
TM/TZV2P	PBE	CCSD(T)	no ^c	0.43	0.87	1.22
TM/TZVP	BP	CCSD(T)	no ^c	0.59	1.21	1.73
TM/TZVP	B3LYP	CCSD(T)	no ^c	0.14	1.39	1.81
ADF/TZP	PBE	CCSD(T)	no ^c	0.66	0.84	1.17
ADF/TZ2P	PBE	CCSD(T)	no ^c	0.75	0.90	1.24
ADF/TZP	BP	CCSD(T)	no ^c	1.04	1.27	1.81

^a MD = mean deviation, MAD = mean absolute deviation, RMSD = root-mean-square deviation. ^b MP2 = dispersion term fitted to reproduce the MP2 curves of the nucleic acid dimers; CCSD(T) = dispersion term fitted to reproduce the CCSD(T) curves of the benzene dimer. ^c Dispersion interaction term for the hydrogen-bonded systems, complexes 1–38, 140, 145–150, and 165 was not included.

interaction between H-bonded fragments reduces slightly the MAD and rmsd values when the PBE functional and a good quality triple ζ basis set is used in connection with the best performing $E_{\text{Disp}}^{\text{CCSD(T)}}$ correction term. Improvement of the performances is quite higher when the dispersion term between H-bonded fragments is omitted with double ζ basis sets. Of course, this is a consequence of the overestimation of H-bond interactions with medium quality basis sets. Using CP corrected interaction energies would probably result in dispersion interaction having a beneficial effect, since it would compensate poorly performing medium quality basis sets. Differently, removing dispersion interaction with the BP and B3LYP functionals deteriorates results. This is a consequence of the different performances of PBE on one side and of BP and B3LYP on the other side, in describing H-bond interactions, a point that will be discussed later.

In conclusion of this part, it is clear that the implementation presented here yields statistical deviations of around 1 kcal/mol from the reference JSCH-2005 benchmark, even if rather small basis sets are used. This is an interesting result, since in the case of biomolecules “real life” systems can easily be hundreds of atoms, and even moderately large triple ζ basis sets, such as the TZVP and TZ2P basis sets, could be impossible to use.

To test further the approach we developed, we calculated the stacking energy between some nucleic base pairs and compared them with those of Spomer et al. CCSD(T) complete basis set

limit data.³⁷ The geometries (see Figure 7) were taken from Spomer’s original work, and the stacking energy was calculated assuming a rigid dissociation, that is, the geometry of each pair of Watson–Crick bonded bases is the same in the stacked pair and in the isolated pair. The calculated interaction energies are displayed in Table 4. These systems already represent a severe test, since dispersion interactions are long-range interactions, and extending the size of the systems can result in unrealistic behavior due to “accumulation” of small contributions from the long attractive tail of the Lennard-Jones potential.⁵⁴ For this test, we used the PBE functional.

First, we note that the uncorrected PBE functional seriously underestimates the stacking energy, with the only exception of the ADF/DZP approach for which fortunate error cancellation leads to an overall rather reasonable agreement with CCSD(T) results. The MP2 parametrizations allow for the reproduction of MP2 data within less than 2 kcal mol⁻¹ using the double ζ basis sets and within less than 1 kcal mol⁻¹ for the triple ζ basis sets. The agreement is a little less favorable for the CCSD(T) parametrizations with a maximum error of 2.5 kcal mol⁻¹ in the case of the CG/GC system with the PBE/ADF/TZP + $\text{Disp}^{\text{CCSD(T)}}$ approach. However, considering the (chemically) crude training data set, the results are promising and show real improvement with respect to uncorrected or Disp^{MP2} corrected PBE.

H-Bonded Base Pairs and Dispersion Interactions. In this section, we discuss the effect of the empirical dispersion interaction term on the energy and geometry of H-bonded base pairs. We start discussing the effect of dispersion interaction on the interaction energy of the H-bonded base pairs in the JSCH-2005 database (complexes 1–38). The statistic data reported in Table 5 clearly indicate that inclusion of dispersion interactions always deteriorates the performance of the uncorrected PBE functional. This effect is reduced with the softer $E_{\text{Disp}}^{\text{CCSD(T)}}$ correction term. Since the interaction energy, in line with similar approaches in this field, is not corrected for BSSE, the stabilizing dispersion term adds to already BSSE overstabilized base pairs. Differently, focusing on the BP and B3LYP functionals, the numbers reported in Table 5 clearly indicate that these two functionals remarkably underestimate the interaction energy of complexes 1–38, and the E_{Disp} correction term compensates this failure, leading to a better agreement with the JSCH-2005 results.

To further investigate this issue, we considered the classical Watson–Crick (WC) G-C and A-T base pairs, see Figure 8. For each base pair (see Table 6), we report the most accurate MP2 interaction energies and geometries taken from ref 38, the CP corrected PBE interaction energies and geometries obtained by us with no E_{Disp} correction term, as well as with the E_{Disp} correction included. Interaction energies which include the BSSE are reported in parentheses.

The first result is the good performance of the PBE functional in reproducing with reasonable accuracy the interaction energy and geometry of H-bonded base pairs when good quality basis sets are used. The BSSE is rather small, and thus both BSSE including and CP corrected values are quite close to the reference value. The good performance with the relatively small SVP basis set is a consequence of a quite higher BSSE,⁴⁵ and the CP corrected values remarkably underestimate the reference values. Inclusion of the E_{Disp} correction term results in a remarkable increase of the interaction energy, roughly 5 kcal/mol both for the G-C and A-T base pairs. Moreover, inclusion

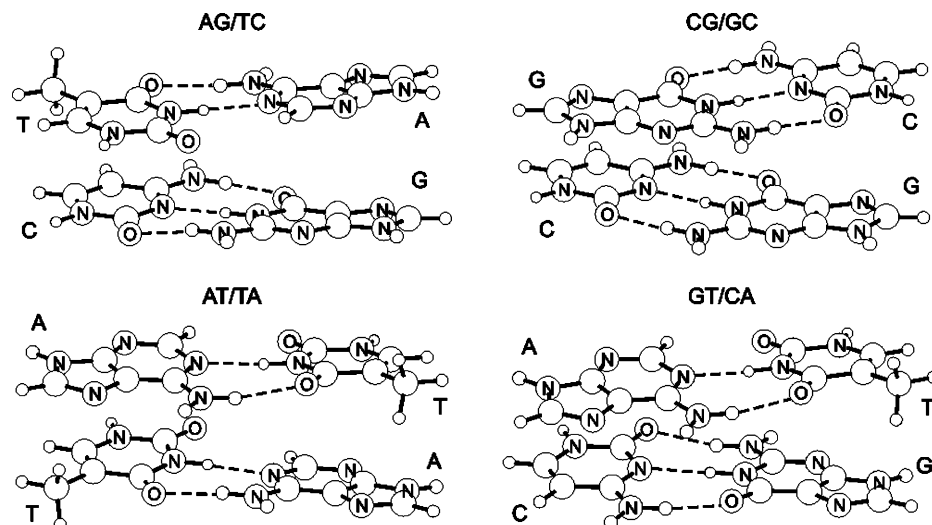


Figure 7. Structures of the different pairs of stacked base pairs.

TABLE 4: Stacked Base Pairs Interaction Energies, E_{stacking} , Calculated with Different Approaches Based on the PBE Functional (Energies in kcal mol⁻¹), and for Comparison, MP2 and CCSD(T) E_{stacking} from Ref 37 Are Also Reported

method	E_{stacking} AG/TC	E_{stacking} AT/TA	E_{stacking} CG/GC	E_{stacking} GT/CA
MP2 ^a	15.3	16.8	19.7	17.1
CCSD(T) ^a	12.5	13.3	17.3	13.4
TM/SVP	4.0	1.8	7.3	1.9
TM/TZVP	-0.9	-2.5	1.8	-3.4
ADF/DZP	9.6	10.0	12.1	9.8
ADF/TZP	-1.5	-2.6	0.6	-4.0
TM/SVP + Disp ^{MP2}	17.5	16.7	21.6	18.1
TM/TZVP + Disp ^{MP2}	16.1	16.3	19.8	17.0
ADF/DZP + Disp ^{MP2}	15.3	16.3	18.1	16.7
ADF/TZP + Disp ^{MP2}	16.3	16.9	19.3	17.2
TM/TZVP + Disp ^{CCSD(T)}	13.9	13.7	17.2	14.0
ADF/TZP + Disp ^{CCSD(T)}	12.2	12.3	14.8	12.1

^a MP2 and CCSD(T) complete basis set limit values from ref 37.**TABLE 5: Statistical Values on Complexes 1–38 (H-Bonded Base Pairs) of the JSCH-2005 Database**

method	MD	MAD	RMSD
PBE/TM/SVP	-3.05	3.05	0.94
PBE/TM/TZVP	1.27	1.33	0.74
BP/TM/TZVP	2.94	2.94	0.77
B3LYP/TM/TZVP	2.80	2.80	0.68
PBE/ADF/DZP	-0.69	0.86	0.91
PBE/ADF/TZP	1.50	1.54	0.69
BP/ADF/TZP	3.13	3.13	0.74
PBE/TM/SVP + Disp ^{MP2}	-6.23	6.23	1.36
PBE/TM/TZVP + Disp ^{MP2}	-2.78	2.78	1.05
PBE/ADF/DZP + Disp ^{MP2}	-2.13	2.13	0.99
PBE/ADF/TZP + Disp ^{MP2}	-2.78	2.78	0.99
PBE/TM/TZVP + Disp ^{CCSD(T)}	-1.71	1.17	0.93
BP/TM/TZVP + Disp ^{CCSD(T)}	-2.30	2.30	1.13
B3LYP/TM/TZVP + Disp ^{CCSD(T)}	-2.64	2.64	1.00
PBE/ADF/TZP + Disp ^{CCSD(T)}	-1.35	1.35	0.82
BP/ADF/TZP + Disp ^{CCSD(T)}	-1.95	1.95	1.04

of the E_{Disp} correction term results in a small shortening, roughly 0.01–0.02 Å, of the H-bonds. Of course, these effects are smaller in the case of the softer $E_{\text{Disp}}^{\text{CCSD(T)}}$ parametrization.

In a previous study, it was reported that inclusion of dispersion interactions in H-bonded systems considerably improve the agreement with high-level reference values.¹² However, the BLYP functional was used, and comparison of the pure DFT calculations with the JSCH-2005 benchmark clearly indicates that the BLYP functional, as the BP and B3LYP

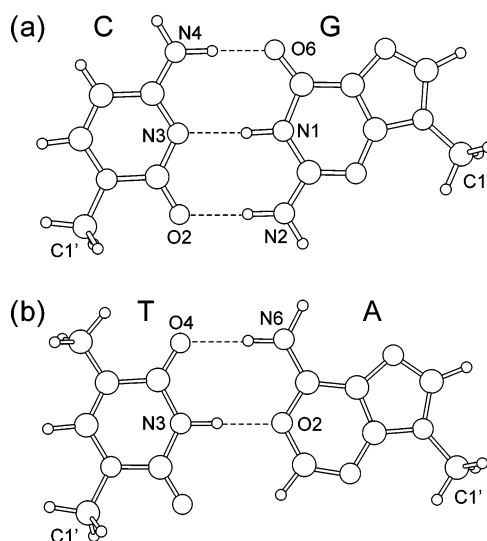


Figure 8. Watson-Crick G:C and A:T base pairs.

functionals, underestimates the stability of H-bonded systems,¹² while the PBE functional results in much better agreement. Thus, it is reasonable that dispersion interactions improve the agreement with reference values when the BLYP, BP, or B3LYP functional are used. Everything depends on the performance of the uncorrected DFT approach. If it performs consistently bad on both H-bonded and dispersion-bonded systems, adding a remarkable dispersion term results in a good agreement with the reference values, and indeed it is a viable solution to the problem. Conversely, if a DFT functional works rather well for a H-bonded system, addition of dispersion interactions degrade performances.¹³

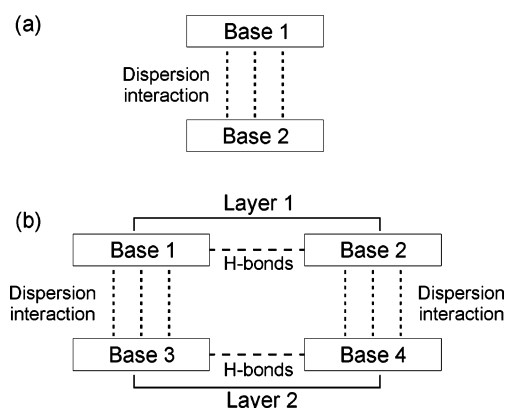
These results suggest that the calculation of H-bonded systems requires either removal of the BSSE or, preferably, removal of the E_{Disp} correction term. In this respect, we note that in many cases it is difficult to decompose the system into two or more fragments as, for example, in the case of a DNA double-helix fragment. Further, BSSE cannot be removed easily during geometry optimization, or when this option is implemented in the program, this requires a remarkable increase of the computational costs. Differently, we believe that it is more convenient to include the E_{Disp} correction term to describe stacking interactions only (for which the DFT–D approach has been originally developed) and to remove it between H-bonded bases.

TABLE 6: CP Corrected PBE Based Interaction Energies (in kcal/mol) and H-Bond Lengths between Heavy Atoms (in Angstroms) for the G-C and A-T Watson–Crick Base Pairs. Energy Values in Parenthesis Include the BSSE^a

G:C pair	ΔE	N2(G)–O2(C)	N1(G)–N3(C)	O6(G)–N4(C)
MP2 ^b	27.5	2.89	2.90	2.75
TM/SVP	22.6 (28.5)	2.87	2.87	2.71
TM/TZVP	26.6 (27.6)	2.90	2.91	2.75
TM/SVP + Disp ^{MP2}	31.2 (32.2)	2.85	2.86	2.70
TM/TZVP + Disp ^{MP2}	31.3 (32.3)	2.87	2.88	2.74
TM/TZVP + Disp ^{CCSD(T)}	29.9 (30.9)	2.90	2.90	2.74

A:T pair	ΔE	N1(A)–N3(T)	N6(A)–O4(T)
MP2 ^b	15.0	2.83	2.86
TM/SVP	9.3 (14.6)	2.78	2.85
TM/TZVP	13.6 (14.4)	2.82	2.89
TM/SVP + Disp ^{MP2}	17.1 (17.9)	2.75	2.84
TM/TZVP + Disp ^{MP2}	17.8 (18.5)	2.77	2.86
TM/TZVP + Disp ^{CCSD(T)}	16.5 (17.3)	2.8	2.88

^a For H-bond labeling see Figure 8. For comparison, the MP2 basis set limit values from ref 38 are also reported. ^b Based on MP2 complete basis set limit values from ref 38.

SCHEME 1

This problem can be solved by flagging the bases (or layers in the case of stacked base pairs as in the DNA double helix) whose interaction needs to be corrected by dispersion interaction. To this end and in analogy with the nonbonded pair-list used by classical molecular dynamics packages, we have included appropriate keywords in both the ADF and the TURBOMOLE implementations to build an atom-based dispersion pair-list that contains the pairs of atoms for which the two-body dispersion term has to be calculated. Thus, in the calculations on stacked base pairs, as in Scheme 1a, atoms belonging to bases 1 and 2 can be flagged differently, and the dispersion pair-list runs over pairs of atoms belonging to base 1 and base 2 only (i.e., atom pairs belonging to the same base are excluded). Similarly, in the calculations on stacked H-bonded base pairs, as in Scheme 1b, atoms belonging to layer 1 (composed by bases 1 and 2) and layer 2 (composed by bases 3 and 4) can be flagged differently, and the dispersion pairs list runs over pairs of atoms belonging to layer 1 and layer 2 only (i.e., atom pairs belonging to the same layer are excluded). That is, the interaction between the H-bonded bases 1 and 2, as well as between bases 3 and 4, is treated at the DFT level only.

Although this is a very simple recipe, we are aware that this is not the ultimate solution, since intermediate situations remain a problem, and it could happen that, during a geometry optimization, a base pair changes its geometry from stacked to H-bonded or vice versa. Nevertheless, in a complex system with well-defined geometry, such as a fragment of DNA double helix or piled G quartets, which present both stacked and H-bonded bases, the rather easy implementation of a dispersion pair-list would allow to correct for dispersion interaction of the stacked

bases only. In this respect, it has to be noted that overestimating the H-bond interaction of a single base pair by 2–3 kcal/mol is not a severe problem. However, if a DNA double-helix fragment composed by 10 base pairs is considered, this error accumulates to the remarkable amount of 20–30 kcal/mol.

Incidentally, we note that in a previous study it was proposed that different parametrizations of the damping function should be used to deal with H-bonded or dispersion-bonded systems, since the former need for a much softer correction.¹³ This could be done easily if the dispersion pair-list we introduced in this paper is used. In fact, within a given system, fragments flagged with different keys (dispersion-bonded fragments) could be treated with standard damping functions and parameters, while fragments flagged with the same key (H-bonded fragments) could be treated with a damping function and parameters properly tuned to cure for missing interaction in H-bonded systems.

As a technical remark, we note that building a dispersion pair-list makes sense when systems presenting both H-bonded and dispersion-bonded fragments are investigated. If only stacked fragments are considered, there is no need to setup a dispersion pair-list.

DFT–D and Solvent Effects. Finally, we tested the DFT–D approach when solvent effects (water) with a continuum solvation model³⁹ (COSMO⁵⁵) are included, and the PBE functional is used. In this case, we considered H-bonded and stacked G-C and A-T base pairs, complexes 1, 3, 117, and 123 of the JSCH-2005 database. Since the JSCH-2005 benchmark is based on gas-phase calculations, we performed single point MP2 calculations on these complexes using the aug-cc-pTVZ basis set both in the gas-phase and in solution using the COSMO implementation of water. The CP corrected MP2 calculations together with the BSSE-including DFT–D calculations at the PBE/TM/TZV2P + Disp^{MP2} level are reported in Table 7. In the case of the DFT–D calculations, we preferred to report BSSE-including values, since this is the standard approach in DFT–D. As already discussed, the numbers reported in Table 7 clearly indicate that in the gas phase the DFT–D approach overestimates the higher-level calculations for the H-bonded geometries, and better agreement can be obtained if dispersion interaction is not included, while the stacked geometries are described better if the E_{Disp} correction term is considered. This is because the reasonably well-performing PBE functional has been used. When solvent effects are included, the numbers reported in Table 7 again indicate that the simple PBE functional reproduces reasonably well the MP2 calculations on the

TABLE 7: MP2 and PBE Based Interaction Energy (in kcal/mol) of H-Bonded (WC) and Stacked (St) G-C and A-T Base Pairs (Complexes 1, 3, 123, and 117 in the JSCH-2005 Database) where MAD Is the Mean Absolute Deviation from the MP2/aug-cc-pVTZ Values

	G-C WC	A-T WC	G-C St	A-T St	MAD
Gas-phase					
MP2/aug-cc-pVTZ	-29.8	-15.8	-8.8	-9.8	
PBE/TZV2P	-29.0	-14.9	0.1	-4.1	4.1
PBE/TZV2P+ Disp ^{MP2}	-33.5	-18.7	-9.3	-12.3	2.4
PBE/TZV2P+ Disp ^{MP2} no-HB ^a	-29.0	-14.9	-9.3	-12.3	1.2
Water					
MP2/aug-cc-pVTZ	-10.9	-7.4	-6.4	-5.2	
PBE/TZV2P	-9.5	-6.5	2.4	3.3	4.9
PBE/TZV2P+ Disp ^{MP2}	-14.0	-10.2	-7.0	-4.9	1.7
PBE/TZV2P+ Disp ^{MP2} no-HB ^a	-9.5	-6.5	-7.0	-4.9	0.8

^a Dispersion correction term is included only for the stacked base pairs.

H-bonded geometries, whereas disagreement from the MP2 calculations is larger when the dispersion-bonded geometries are considered. Inclusion of dispersion interactions deteriorates the results for the H-bonded geometries, whereas it cures very nicely the results for the dispersion-bonded geometries. Again, lower deviation from the reference values is found if dispersion interaction is included only for dispersion-bonded geometries. In conclusions, these tests indicate that, very promisingly, the DFT-D approach can perform well even if the solvent-independent and gas-phase parametrized E_{Disp} term is added to in solvent calculations, although it is clear that more severe testing has to be done.

Conclusions

We have presented an accurate parametrization of density functional theory including an empirical correction term to describe properly π -stacking interactions in nucleic acids. Specifically, we propose a variety of parametrizations to account for dispersion interactions when using the PBE, BP, and B3LYP functionals in combination with the ADF DZP and TZP basis sets and Turbomole SVP and TZVP basis sets. Of course, parametrization for the SVP and TZVP basis sets can be used with the PBE, BP, and B3LYP implementation in any quantum mechanics code based on Gaussian-type basis sets.

The developed parameters have been tested versus the JSCH-2005 benchmark, with a rmsd usually comprised between 1 and 2 kcal/mol. Better performances are more often achieved using parameters from parametrization against CCSD(T) benzene dimer interaction energies, than from parametrization versus MP2 interaction energies between stacked nucleic acid base pairs. This is interesting, since the training set based on the CCSD(T) benzene dimer interaction energies is rather limited and the benzene dimer is quite different from nucleic base pairs. Additionally, we tested our approach versus extended systems composed by stacked H-bonded base pairs, and we also tested the performances of the model when solvent effects (water) are included in the calculations. Also, in these cases, the approach we presented here is able to reproduce high level calculations satisfactorily.

Further, we have shown that accurate definition of fragments that need to be corrected for dispersion interaction can also improve agreement with the JSCH-2005 benchmark. Indeed, inclusion of dispersion interactions between H-bonded base pairs results in a slight overestimation of the interaction energy when the PBE functional is used. Although the effect on geometries

and energies is not dramatic, in our approach this problem can be avoided through the easy definition of a dispersion pair-list. Differently, because of the weak performances of the BP and B3LYP functionals on H-bonded systems, inclusion of the dispersion correction leads to a better agreement with the JSCH-2005 benchmark.

Although we focused on an accurate parametrization for nucleic acids, because of their remarkable relevance, the method we used to parametrize properly the damping function is rather general. A similar approach could be applied for the accurate parametrization of other functionals, basis sets, differently derived and combined C_6 parameters, but also to molecular systems different from nucleic acids. Since we have shown that the parametrization is dependent on both the basis set and the functional, using the same damping function with different basis sets and/or functionals could lead to inconsistent results. In this respect, the damping functions we presented could be valid for studying other kind of van der Waals complexes, provided that appropriate test calculations are performed, since the nature of the interaction remains essentially the same.

Finally, we focused our study on nucleic acids base stacking because of their biological relevance. The correct reproduction of structure and energetic of nucleic acids requires the development of methods able to treat stacking effects properly, since the other two main contributions to structure and energetic of nucleic acids systems (H-bonds and solvent effects) are already evaluated with great accuracy.

Acknowledgment. We thank one of the reviewers for useful suggestions. This work was supported by CINECA (Progetto Supercalcolo CINECA/INSTM).

Supporting Information Available: DFT-D interaction energies of the JSCH-2005 benchmark. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Hobza, P.; Sponer, J.; Polasek, M. *J. Am. Chem. Soc.* **1995**, *117*, 792.
- (2) Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331.
- (3) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143.
- (4) Zimmerli, U.; Parrinello, M.; Koumoutsakos, P. *J. Chem. Phys.* **2004**, *120*, 2693.
- (5) Dobson, J. F. *Int. J. Quantum Chem.* **1998**, *69*, 615.
- (6) Kohn, W.; Meir, Y.; Makarov, D. E. *Phys. Rev. Lett.* **1998**, *76*, 1780.
- (7) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.
- (8) Ahlrichs, R.; Penco, R.; Scoles, G. *Chem. Phys.* **1977**, *19*, 119.
- (9) Mooij, W. T. M.; van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Eijck, B. P. *J. Phys. Chem. A* **1999**, *103*, 9872.
- (10) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149.
- (11) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515.
- (12) Morgado, C.; Vincent, M. A.; Hillier, I. H.; Shan, X. *Phys. Chem. Chem. Phys.* **2007**, *9*, 448.
- (13) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. *J. Comput. Chem.* **2007**, *28*, 555.
- (14) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287.
- (15) Piacenza, M.; Grimme, S. *Chem. Phys. Chem.* **2005**, *6*, 1554.
- (16) Piacenza, M.; Grimme, S. *J. Am. Chem. Soc.* **2005**, *127*, 14841.
- (17) Reha, D.; Kabelac, M.; Ryjacek, F.; Sponer, J.; Sponer, J. E.; Elstner, M.; Suhai, S.; Hobza, P. *J. Am. Chem. Soc.* **2002**, *124*, 3366.
- (18) Otyepka, M.; Sklenovsky, P.; Horinek, D.; Kubar, T.; Hobza, P. *J. Phys. Chem. B* **2006**, *110*, 4423.
- (19) Oliva, R.; Tramontano, A.; Cavallo, L. *RNA* **2007**, *13*, 1427.
- (20) Grimme, S.; Diedrich, C.; Korth, M. *Angew. Chem., Int. Ed.* **2006**, *45*, 625.
- (21) Du, A. J.; Smith, S. C. *Nanotechnology* **2005**, *16*, 2118.
- (22) Feng, S.; Li, T. *J. Chem. Theory. Comp.* **2006**, *2*, 149.
- (23) Neumann, M. A.; Perrin, M.-A. *J. Phys. Chem. B* **2005**, *109*, 15531.

- (24) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (25) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- (26) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.
- (27) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
- (28) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (29) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (30) Ahlrichs, R.; B  r, M.; H  ser, M.; Horn, H.; K  lmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.
- (31) Sch  fer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.
- (32) Sch  fer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.
- (33) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. *Theor. Chem. Acc.* **1998**, *99*, 391.
- (34) te Velde, G.; Bickelhaupt, F. M.; van Gisbergen, S. J. A.; Fonseca Guerra, C.; Baerends, E. J.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931.
- (35) *ADF 2005.01, SCM, Theoretical Chemistry*; Vrije Universiteit: Amsterdam, The Netherlands; <http://www.scm.com>.
- (36) Jure  ka, P.; Sponer, J.; Cern  y, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.
- (37) Sponer, J.; Jure  ka, P.; Marchan, I.; Luque, F. J.; Orozco, M.; Hobza, P. *Chem. Eur. J.* **2006**, *12*, 2854.
- (38) Sponer, J.; Jure  ka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142.
- (39) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.
- (40) Slater, J. C.; Kirkwood, J. G. *Phys. Rev.* **1931**, *37*, 682.
- (41) Buckingham, A. D.; Fowler, P. W.; Hutson, J. M. *Chem. Rev.* **1988**, *88*, 963.
- (42) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533.
- (43) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827.
- (44) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.
- (45) Muller, A.; Losada, M.; Leutwyler, S. *J. Phys. Chem. A* **2004**, *108*, 157.
- (46) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- (47) Kendall, R. A.; Dunning, T. H. J.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- (48) Weigend, F. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285.
- (49) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- (50) Treutler, O.; Ahlrichs, R. *J. Chem. Phys.* **1995**, *102*, 346.
- (51) Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2004**, *6*, 5119.
- (52) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.
- (53) Gauss, J.; Stanton, J. F. *J. Phys. Chem. A* **2000**, *104*, 2865.
- (54) Cavallo, L.; Woo, T. K.; Ziegler, T. *Can. J. Chem.* **1998**, *76*, 1457.
- (55) Klamt, A.; Schueuermann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799.