Density functional theory with dispersion corrections for supramolecular structures, aggregates, and complexes of (bio)organic molecules

Stefan Grimme,* Jens Antony, Tobias Schwabe and Christian Mück-Lichtenfeld

Received 20th October 2006
First published as an Advance Article on the web 26th January 2007
DOI: 10.1039/b615319b

Kohn–Sham density functional theory (KS-DFT) is nowadays the most widely used quantum chemical method for electronic structure calculations in chemistry and physics. Its further application in *e.g.* supramolecular chemistry or biochemistry has mainly been hampered by the inability of almost all current density functionals to describe the ubiquitous attractive long-range van der Waals (dispersion) interactions. We review here methods to overcome this defect, and describe in detail a very successful correction that is based on damped $-C_6 \cdot R^{-6}$ potentials (DFT-D). As examples we consider the non-covalent inter- and intra-molecular interactions in unsaturated organic molecules (so-called π–π stacking in benzenes and dyes), in biologically relevant systems (nucleic acid bases/pairs, proteins, and 'folding' models), between fluorinated molecules, between curved aromatics (corannulene and carbon nanotubes) and small molecules, and for the encapsulation of methane in water clusters. In selected cases we partition the interaction energies into the most relevant contributions from exchange-repulsion, electrostatics, and dispersion in order to provide qualitative insight into the binding character.

1 Introduction

Non-covalent interactions are playing an increasingly important role in modern chemical research and are considered as cornerstones in supramolecular chemistry, materials science, and even biochemistry.[1–6] Although a very detailed understanding on an atomic or molecular level is still lacking, important progress has been achieved in recent years in the quantum-mechanical description of the relevant forces.[7] Nowadays, the accurate computation of structures and interaction potentials for small molecular complexes (<10 atoms) with wave-function-based methods like second-order Møller–Plesset perturbation theory (MP2)[8,9] in combination with coupled-cluster methods (*e.g.* CCSD(T)[10]) has become possible (see *e.g.* the reviews by Sherrill,[11] Tzusuki,[12] or Hobza[13] *et al.*). Using these *ab initio* techniques, even notoriously difficult systems with dominant van der Waals (vdW, dispersion) contributions such as dimers of aromatic molecules (of so-called π–π stacked or CH–π type, for overviews see refs. 2,14) can be investigated. Dispersion interactions are ubiquitous, long-range attractive forces which act between separated molecules even in the absence of charges or permanent electric moments. They stem from many-particle (electron correlation) effects that are complicated by the quantum-mechanical wave-nature of matter.[7,10,15]

The interactions between aromatic groups in DNA and RNA and between aromatic side chains in proteins are significantly influenced by vdW forces.[16–18] A further improved understanding of non-covalent interactions and an even more accurate description would greatly aid the rational synthesis of functional supramolecular structures or drug design, *e.g.* for anti-cancer agents that intercalate into DNA, also in a π–π stacking mode.[19] Currently, the benzene dimer as a model system is being extensively investigated in this context, and at least four high-quality *ab initio* studies have appeared within the last year.[20–23]

However, accurate *ab initio*-correlated wavefunction methods are computationally too demanding for routine studies of the non-covalent interactions in larger (>50 atoms), chemically interesting systems. Using special techniques, perturbation methods can be used nowadays for complexes of about 100 atoms, as recently shown for a C_{60}–tetraphenylporphyrin single point calculation.[24] In particular, there is great need for effcient and robust quantum-mechanical approaches that allow structure optimization. Such structural information usually forms the basis for a deeper understanding of the system's functionality but is often difficult to obtain experimentally.

Kohn–Sham density functional theory[25] (KS-DFT) is now the most widely used method for electronic structure calculations in condensed matter physics and quantum chemistry.[26,27] This success mainly results from significant 'robustness', *i.e.*, providing reasonably accurate predictions for many properties of various molecules and solids at affordable computational expense. However, a general drawback of all common density functionals, including hybrids like the popular B3LYP, is that they can not describe long-range electron correlations that are responsible for the dispersion forces.[28–30] The DFT problem for vdW interactions now has become a very active field of research and therefore, we will provide a summary of the most recent approaches in Section 2.2. From the practical point of view, where the focus is on robustness and computational speed, empirical $-C_6 \cdot R^{-6}$ corrections to standard density functionals seem most promising. The most widely applied and very well-tested approach is DFT-D,[31,32] which provides high accuracy in many different situations.[33–37]

A review and general perspective of the DFT-D method in a variety of different applications is the main aim of this paper, which is organized as follows. In Section 2 we present a brief and qualitative description of the theory of non-covalent interactions. We explain in some detail the problems of KS-DFT with vdW

Theoretische Organische Chemie, Organisch-Chemisches Institut der Universität Münster, Corrensstraße 40, D-48149, Münster, Germany. E-mail: grimmes@uni-muenster.de; Tel: +49-251-8336512

Stefan Grimme studied Chemistry at the TU Braunschweig and finished his PhD in 1991 in Physical Chemistry on laser spectroscopy. He then moved to Bonn where did his Habilitation in Theoretical Chemistry in the group of Sigrid Peyerimoff. In 2000 he got the C4 chair for Theoretical Organic Chemistry at the University of Münster. His research interests are the development of quantum chemical methods for large systems, density functional theory, electronic spectroscopy/excited states, and the properties of chiral systems.

Jens Antony earned his PhD in computational chemistry at the University Lübeck under the supervision of Professor Alfred Trautwein. After postdoctoral stays at UC Davis with Professor Alexei Stuchebrukhov, the Veterinary and Agricultural University of Denmark with Professor Rogert Bauer, and the FU Berlin with Professor Christof Schütte, he joined the group of Professor Grimme at the University Münster in 2006. His main research activity is the application of quantum-chemical and molecular-dynamics methods to systems of biological interest.

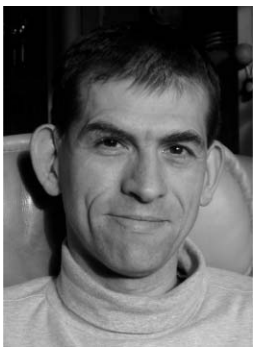
Tobias Schwabe studied chemistry at the University of Münster. In his advanced courses he focused on theoretical chemistry. He received his diploma degree for his thesis on the B2-PLYP and the mPW2-PLYP functionals. He is now researching for his PhD in DFT development under the supervision of Professor Grimme.

Christian Mück-Lichtenfeld finished his PhD on the mechanism and stereoselectivity of carbolithation reactions with Professor Hubertus Ahlbrecht at Justus-Liebig-University in Giessen in 1998. He then started a post-doctorate in computational chemistry at the University of Münster, working with Professor Ernst-Ulrich Würthwein until 2000. After that, he completed the change to theoretical chemistry and joined the Grimme group. He is now senior lecturer, and his research interests include organic/organometallic reaction mechanisms and non-bonding interactions.

Stefan Grimme

Jens Antony

Tobias Schwabe



Christian Mück-Lichtenfeld

interactions and outline possible solutions to the problem. After a short review of the DFT-D method in Section 2.4 we discuss some more technical aspects of the computations that are necessary to follow and understand this paper. The Results and discussion section is split into four parts. First, we decompose the interaction energies for model complexes into physically meaningful components and show that the empirical dispersion correction used has a very solid physical basis. We then compare DFT-D results for a set of vdW and hydrogen-bonded complexes to coupled-cluster reference data in order to get some impression about the accuracy of the methods. The chemical examples are grouped into intra- and intermolecular cases and have been selected to cover a broad range of problems in organic or supramolecular chemistry. We consider π-stacking in triptycenes and in the anthracene dimer, folding of alkane chains, complexes of small molecules with the bowl-shaped corannulene and carbon nanotubes, dimers of organic dyes, and the formation of methane hydrates.

2 Theory

2.1 General

Although there is no rigorous quantum-mechanical definition of non-covalent interactions, there is common understanding about an operational definition.[7] The relevant molecular or atomic fragments are separated by distances where the overlap of one-electron functions (orbitals) is so small that the covalent (quantum-mechanical interference) or charge-transfer character of bonding is negligible. There are of course borderline cases where the classification (non-covalent *vs.* covalent/ionic) strongly depends on the method used for analysis, but all examples presented here are uncritical in that respect (for a more detailed discussion about bonding in such problematic cases the reader is referred to the recent review of Popelier[38]).

The term non-covalent interaction (sometimes also misleadingly called non-bonded in the force-field community) is usually associated with the formation of weakly bonded complexes, aggregates, or even condensed phases from separable (individual) units A and B (intermolecular case, see top of Fig. 1).

This situation is also theoretically most simple because the interaction energy ΔE can easily be computed from the complex and fragment total energies (supermolecular approach)

$$\Delta E = E(\mathrm{AB}) - E(\mathrm{A}) - E(\mathrm{B}). \quad (1)$$

A perturbational view can also be appropriate in this case, which leads to a more detailed picture of binding. If two separate

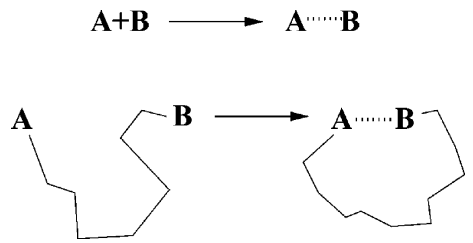


Fig. 1 Non-covalent interaction modes between molecular units or fragments A and B in the inter- (top) and intramolecular case (bottom).

Hamilton operators for fragments A and B and a Coulombic interaction term $\hat{V}$ as a perturbation are defined,

$$\hat{H} = \hat{H}(\mathrm{A}) + \hat{H}(\mathrm{B}) + \hat{V}(\mathrm{AB}), \quad (2)$$

and the Pauli-principle (anti-symmetry with respect to inter-electronic permutations) is obeyed, the resulting method is called symmetry-adapted perturbation theory (SAPT).[39] SAPT results are considered in more detail in Section 3.1. Note, however, that a very important class of non-covalent interactions occurs in an intramolecular fashion (Fig. 1, bottom), *e.g.* base-stacking in DNA/RNA or protein folding. It is quite clear that the interaction between fragments or groups A and B is not fundamentally different if they form a complex from two separated entities or if they are *e.g.* connected by a flexible and chemically inert alkane chain. In practice it seems highly desirable to have theoretical methods that treat inter- and intramolecular cases on an equal footing. For example, SAPT does not belong to this class because $\hat{V}$ can not be regarded as a perturbation in the intramolecular case (and also the popular counterpoise correction is not defined; see Section 2.4).

Nevertheless, SAPT is very useful as it provides a solid theoretical basis for qualitative understanding and classification of non-covalent bonding. In the simplest picture, one can distinguish three dominant types of non-covalent interactions. The exchange repulsion (EXR) term stems from the Pauli or anti-symmetry principle, *i.e.*, electrons with the same spin can not occupy the same region of space. This can be seen in a simple orbital picture (Fig. 2). The anti-bonding linear combination of the fragment orbitals (upper level) is always more destabilizing than the bonding orbital stabilizing, so that in total a repulsive energy contribution always remains between any closed shells. Because the EXR mainly involves the overlap between fragment orbitals, it is short-ranged and decays exponentially with distance. The EXR is quite accurately described by all theoretical models that treat orbital overlap exactly, *i.e.*, by Hartree–Fock (HF) and KS-DFT but not by most semiempirical models (*e.g.* INDO, MNDO, AM1$\cdots$) that absorb EXR empirically into core–core repulsion functions.[40]

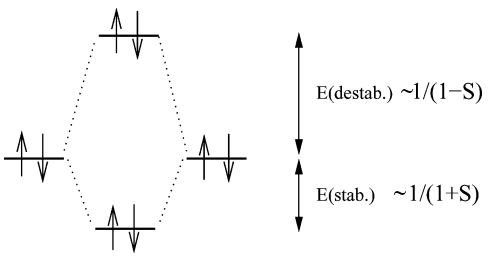


Fig. 2 Closed-shell orbital interaction.

The second part is due to electrostatic interactions (ES). These can be further decomposed into a purely static (first-order) term that arises from the unperturbed interactions of the charge distributions of the fragments, and an induction (second-order) term that results from polarization of one charge distribution by the electric moments of the other fragment and subsequent interaction. The ES terms can be repulsive or attractive depending strongly on the inter-fragment orientation (*e.g.* parallel or anti-parallel dipole moments) and are usually long-range. At short distances and for not very polar molecules, however, the ES contributions are often attractive because the nuclear charges are not fully screened by the electrons and thus, they can attract the electron density from the other fragment. The interaction potentials resulting from the ES part are accurately described by all methods that yield good electron densities (molecular multipole-moments) and reasonable electric polarizabilities. KS-DFT methods perform in this respect better than HF (which yields molecules that are too polar and too unpolarizable). This is the reason to use a KS-DFT description of the monomers in SAPT (called SAPT-DFT[41–43]). An accurate description of the ES effects is a particular problem for classical force fields that often employ a crude atomic point-charge model for this purpose and mostly neglect induction (for the development of polarizable force-fields see *e.g.* ref. 44).

Last but not least we must consider the weak but ubiquitous attractions between electron clouds that do *not* significantly overlap, arising from instantaneous transition dipoles (where an electron transiently fluctuates from a filled level to an empty level) on each system that couple to each other. These dispersion (van der Waals, vdW) interactions are a quantum-mechanical electron correlation effect that is completely absent when considering the classical interactions of separate charge distributions. Their strength can readily be shown[7,15,45] to decrease as R^{-6} with the separation, R, of the two systems in the non-overlapping (asymptotic) regime. When the charge distributions overlap, electron correlation effects continue to increase in strength, but are then usually smaller than strong orbital repulsions between filled levels (EXR) that prevent the systems from approaching closer than their typical vdW radius. The accurate account of dispersion effects is most difficult in the quantum-mechanical treatment of non-covalent interactions and for weakly bonded complexes in particular. Because the dispersion terms are often of the same size as but of different sign (attractive) than the EXR, their neglect or insufficient treatment usually leads to binding that is too weak, or even to no binding at all. Turning it the other way around, one can define vdW complexes by their property of being unbound at the (uncorrelated) HF level of theory.

Because of their non-local (long-range) character, the dispersion interactions are accurately accounted for only by correlated wavefunction methods like MP2, CCSD(T), or by fully correlated methods like quantum-Monte-Carlo,[20,46] but constitute a serious problem for KS-DFT. This aspect of DFT is considered in more detail in the next section.

2.2 Density-functional-based methods

More than a decade it has been known that the commonly used density functionals do not describe dispersion interactions correctly.[28–30,47–51] In the following we first want to discuss the

problems of KS-DFT with the vdW interactions qualitatively, as schematically outlined in Fig. 3. We then give an overview of the existing approaches in the literature that try to overcome these defects. Note that we are just talking about the correlated part of the interaction which is more-or-less dispersion; all the other interactions mentioned in the preceding section are of course accounted for by KS-DFT.

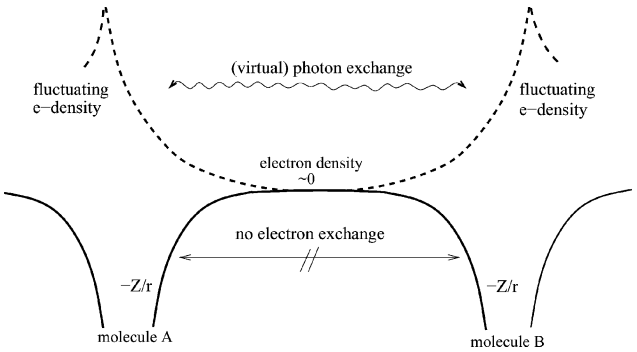


Fig. 3 Sketch of electron densities (dashed lines) resulting from molecular Coulombic potentials (solid lines) for large intermolecular distances and the possible interaction modes (exchange of virtual photons leading to dispersion forces and electron exchange leading to charge-transfer and covalent interactions). Adapted with small modifications from a recent talk by P. Pulay.

Due to the Coulomb potential of the nuclei in molecules, the resulting electron densities (or equivalently the corresponding wavefunctions) exponentially decay with respect to distance from *e.g.* the center of their electronic charge. When molecules (or groups) start to interact but are separated by large distances (*e.g.* >5–10 Å), the electron density between them is approximately the sum of the fragment densities and is vanishingly small. This is illustrated in Fig. 3 by the broken arrow that indicates that non-classical effects like electron exchange (leading to charge-transfer and covalent interactions) is then more-or-less forbidden due to the very broad (and also high) 'tunnelling' barrier between the fragments.

More precisely, however, the electron density of the supermolecule is in fact very slightly different from the sum (in all regions of space), and thus with the (unknown) exact density functional one can in principle (in agreement with the first Hohenberg–Kohn theorem[52]) also compute the correlated part of the interaction energy. With approximate density functionals, however, this turns out to be extremely difficult because the densities and their changes are very, very small and furthermore not very specific with respect to these correlation effects. Fig. 3 also illustrates that the physics of the dispersion interaction should be described by exchange of virtual photons stimulated by the fluctuating (correlated) electrons. By some means, DFT must also account for this phenomenon by allowing 'excitations' in the fragments. This, however, is out of the range of standard KS-DFT, which does not contain any information about the necessary virtual orbitals.

Most attempts to correct for the dispersion problem in DFT can be classified into three groups:

1. *Conventional functionals in the generalized gradient approximation (GGAs) including hybrids or meta-hybrids.* Because in real complexes there often remains a small electron density between the fragments, adjusted, specially selected or designed density functionals may account for some dispersion effects. In the last two years, several such functionals have been proposed.[53–57] Note that these methods usually provide quite large errors for pure vdW complexes and may often not work equally well for all types of non-covalent interactions (*e.g.* they may not provide a consistent picture of stacking *vs.* H-bonding in nucleic acid (NA)–base pairs).

2. *Special correlation functionals or orbital-based DFT methods.* These methods try to incorporate the basic physics of dispersion *via* non-local, orbital-dependent ansatzes.[53,58–72] Most of these approaches are at a quite 'experimental' stage and are furthermore computationally more costly than standard KS-DFT.

3. *DFT/molecular mechanics (MM)-based hybrid schemes.* The researchers in this group try to circumvent the inherent problems with the electron density by empirical dispersion corrections of the form $-C_6 \cdot R^{-6}$ added to existing density functionals.[31,73–79] These methods focus on a realistic description of the asymptotics of the problem. They have been applied successfully in various contexts.[33–36,80–82] The idea originally emerged from Hartree–Fock calculations.[83–86] Our ansatz, called DFT-D,[31,32] is also of this type and is described in more detail in Section 2.3.

In the following, the most important and representative examples from each of the above-mentioned groups are listed.

Group 1. Becke's 'half-and-half" functional BH&H reproduces binding energies and potential energy surfaces for π-stacked geometries of substituted benzenes and pyridines, as well as pyrimidine and DNA bases[53] within ±0.5 kcal mol^{-1} of MP2 and/or CCSD(T) reference data. This result is presumably due to error cancellation (as the authors state), and hydrogen bonding interaction energies are significantly overestimated with this functional. For vdW complexes of fluorine-containing organic molecules, only the PBE density functional yields some binding that is, however, much too weak compared to the quite accurate MP2 results.[87] The X3LYP functional improves the accuracy of hybrid GGA methods for rare-gas dimers[54] and the water dimer[88] significantly, but fails qualitatively for stacking[47] which contradicts the original claims of its inventors that it is well-suited for non-bonded interactions. The Wilson–Levy correlation functional together with Hartree–Fock exchange reproduces binding trends for selected rare-gas dimers, isomers of the methane dimer, benzene dimer, naphthalene dimer, and stacked base-pair structures.[55]

Some meta-GGAs incorporating kinetic energy density have been assessed to quantitatively account for dispersion effects. Zhao and Truhlar[89] describe a test of 18 density functionals for the calculation of bond lengths and binding energies of rare gas-dimers, alkaline-earth metal dimers, zinc-dimer, and zinc–rare-gas dimers. The authors conclude, from the combined mean percentage unsigned error in geometries and energies, that M05-2X[56] and MPWB1K[57] are the overall best methods for the prediction of vdW interactions in the 17 metal and rare-gas vdW dimers studied. A set of 13 complexes with biological relevance is considered in a study of newly developed DFT methods, which give reasonable results for the stacked arrangements in the DNA base pairs and amino acid pairs, in contrast to previous DFT methods, which fail to describe interactions in stacked complexes.[90] In ref. 91 multi-coefficient extrapolated density functional theory methods are used to calculate the interaction energy of benzene dimers. The TPSS and TPSSh functionals[92] produce vdW bonds in ten rare-gas dimers with $Z \leq 36$ and correct the overbinding

of the local spin density approximation.[93] Binding energy curves for the ground-state rare-gas diatomics Ne_2 and Ar_2 and for the alkaline-earth diatomic Be_2 in reasonable agreement with those from experiment are found for PBE and TPSS, but they have an incorrect asymptotic behavior for large internuclear separations.[94] Adamo and Barone[95] achieved an improved description of He_2 and Ne_2 interaction potentials, with accurate results for other properties too, by refitting the Perdew–Wang exchange functional and using it in a hybrid model called mPW1PW.

Group 2. Langreth and Lundqvist developed a nonlocal correlation energy functional for dispersion interactions.[68,70–72,96,97] They applied it successfully to graphitic systems, polycyclic aromatic hydrocarbon dimers,[98] parallel polymers,[99] and benzene dimers.[61,62] In ref. 65, vdW interactions in the He and Ne dimers are modelled within a local-orbital DFT formulation, which is stated to be suitable for generalization to the case of weakly interacting large systems.

Range-separated schemes treat electron interactions at short distances by DFT, while for long distances wavefunction methods are applied, thereby correcting the incorrect asymptotic behavior of exchange functionals derived from approximate density functionals.[64] Potential curves for alkaline-earth dimers with a range separated hybrid method with perturbational long-range correlation corrections offer a dramatic improvement over conventional DFT approaches.[63] A short-range gradient-corrected density functional combined with a long-range coupled-cluster scheme has been applied to all homo- end heteronuclear rare-gas dimers of He, Ne, Ar, Kr, and Xe atoms.[100] The long-range-corrected density-functional theory with the Andersson–Langreth–Lundqvist vdW functional was applied to the calculation of the π-aromatic interaction of the benzene and naphthalene dimers.[67]

Group 3. Elstner *et al.* extended their self-consistent-charge, density-functional tight-binding (SCC-DFTB) method by a damped London-type dispersion formula.[73] The C_6 coefficients are derived from atomic polarizabilities. Compared to the unaugmented SCC-DFTB method, slight improvement is obtained in case of H-bonded base pairs, while the performance is changed qualitatively in case of stacked DNA base pairs relative to MP2 data by providing quite accurate interactions that are not present originally. In ref. 101, a more extensive testing of the procedure has been recommended. Wu and Yang[75] derived the atomic C_6 coefficients from C_6 coefficients for molecule pairs obtained from dipole oscillator strength distributions. Using four density functionals and two forms of the damping function, the method is applied to rare-gas diatomic molecules, stacking of base pairs, and the conformational stability of polyalanines, with results comparable to MP2 results in many cases. Scoles and co-workers[74] use a hybrid approach in which the dispersion energy is obtained by a damped multipolar expansion to calculate interaction energies and structural parameters of the rare gas dimers Ar_2 and Kr_2, the water dimer, the benzene dimer, and three metal carbonyls. The dispersion coefficients are derived from the polarizabilities and ionization potentials of the interacting molecules. Zimmerli *et al.*[76] compare the performance of the correction terms and accompanying damping functions of the three aforementioned approaches[73–75] in combination with different exchange correlation functionals by application to water–benzene dimer geometries.

Becke and Johnson used a similar method but calculated the C_6 dispersion coefficients specifically for the system under investigation from the dipole moment of the exchange-hole,[77,78] and distributed these coefficients between the atoms. The approach also uses a different damping function and was subsequently extended to include C_8 and C_{10} coefficients as well.[79] Calculations with this method and a specially chosen functional for 45 vdW complexes resulted in remarkably accurate inter-molecular separations and binding energies compared to high-level reference data.[81,82]

The method of Lilienfeld *et al.*[102–104] is a bit different as it uses optimized, atom-centered non-local potentials that are normally used in the context of pseudopotentials for core-electrons. The application of this method for modelling attractive long-range vdW forces is illustrated for argon–argon, benzene–benzene, graphite–graphite, argon–benzene, Ar_nKr_m ($n + m \leq 4$) vdW clusters, cyclooctatetraene, and the hydrogen bromide dimer $(HBr)_2$. The approach has also been applied to calculate interaction energies for polyaromatic hydrocarbon molecules from monocyclic benzene up to hexabenzocoronene[105,106] and the adsorption of Ar on graphite.[107] In contrast to the $-C_6 \cdot R^{-6}$ approaches, this dispersion correction also produces changes in the electronic charge density. A serious disadvantage is, however, that the potentials do not show the correct asymptotic R^{-6} behavior and decay too fast (exponentially) with interatomic distance.

An MP2/DFT hybrid method to study both bond-rearrangements and vdW interactions is proposed by Tuma and Sauer.[108–110] The embedding scheme combines, similar to the ONIOM method,[111] MP2 calculations for the reaction site with DFT calculations for a large extended system, which are extrapolated to the complete basis set limit and the full periodic structure, respectively. The approach has been applied to the protonation of isobutene in zeolites.

2.3 The DFT-D approach

The idea to treat the difficult dispersion interactions classically and to combine the resulting potential with a quantum chemical approach (a kind of QM/MM scheme) goes back to the 1970s in the context of HF theory[83,84] (for more recent HF + disp models see ref. 112,113). The method has been forgotten for almost 30 years and was rediscovered a few years ago as the DFT problems became more evident.[31,73,75–79] The basic idea of our work was to develop a robust dispersion correction that can be applied (without any laborious highly specific fitting procedure) to common standard density functionals. Concomitantly, the approach should be as simple as possible, and in particular allow the easy calculation of energy gradients for efficient geometry optimization, which is one of the main purposes of the method.

For the dispersion correction in the KS-DFT formalism we use a slightly modified version[32] of the approach described originally in ref. 31. The total energy is given by

$$E_{\mathrm{DFT\text{-}D}} = E_{\mathrm{KS\text{-}DFT}} + E_{\mathrm{disp}}, \tag{3}$$

where $E_{\mathrm{KS\text{-}DFT}}$ is the usual self-consistent Kohn–Sham energy as obtained from the chosen density functional and E_{disp} is an empirical dispersion correction given by

$$E_{\mathrm{disp}} = -s_6 \sum_{i=1}^{N_{\mathrm{at}}-1} \sum_{j=i+1}^{N_{\mathrm{at}}} \frac{C_6^{ij}}{R_{ij}^6} f_{\mathrm{dmp}}(R_{ij}). \tag{4}$$

Here, N_{at} is the number of atoms in the system, C_6^{ij} denotes the dispersion coefficient for atom pair ij, s_6 is a global scaling factor that only depends on the DF used, and R_{ij} is an interatomic distance. In order to avoid near-singularities for small R and electron correlation double-counting effects, a damping function f_{dmp} must be used, which is given by

$$f_{dmp}(R_{ij}) = \frac{1}{1 + e^{-d(R_{ij}/R_r - 1)}}, \quad (5)$$

where R_r is the sum of atomic vdW radii in the ground state. We recently determined these radii (taken as the radius of the 0.01 a_0^{-3} electron density contour from atomic ROHF/TZV computations) and atomic C_6 coefficients for the elements H–Xe. For the pairwise dispersion coefficients we employ a simple geometric mean of the form

$$C_6^{ij} = \sqrt{C_6^i C_6^j}. \quad (6)$$

The interatomic dispersion potential is shown for the two examples carbon and fluorine in Fig. 4. As can be seen, the damping function effectively reduces the interactions to zero typically below 200 pm. Asymptotically for large R it approaches unity, and thus the required R^{-6} dependence is recovered. At intermediate distances we get a minimum in the potential that lies usually slightly below the sum of the corresponding atomic vdW radii, *e.g.* at about 300 pm for carbon. The potential for the two fluorine atoms looks qualitatively similar but as the atom is smaller and less polarizable, the interaction potential is overall smaller and the minimum is found at shorter distances.

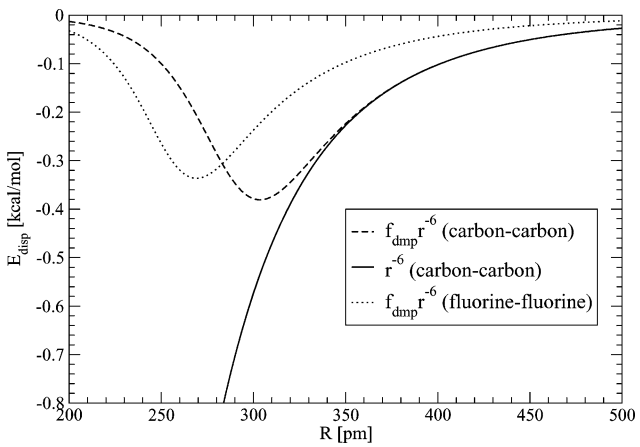


Fig. 4 Dispersion correction (eqn (4)) for two carbon and two fluorine atoms ($s_6 = 1.0$) separated by a distance R. The solid line shows the undamped potential for comparison.

2.4 Technical aspects

Non-covalent interactions are typically one to two orders of magnitude smaller per atom pair than covalent or ionic interactions. Furthermore, the interaction energies are tiny compared to the total electronic energies of the complex and the fragments that are used in the super-molecular calculation of the interaction energy. This has been illustrated with the analogy of attempting to 'weigh the captain of a ship by weighing the ship with and without the captain'. Because the individual energies are necessarily computed by approximate methods, their shortcomings may then show up as large errors in computed potentials. In practice we must rely on the fact that the errors for the individual energy computations ('weighings') are systematic and cancel out when taking the difference.

When using wavefunction-based methods, one usually considers one-electron (AO basis set incompleteness) and many-particle errors (electron correlation treatment). Today's standard procedure is to perform a series of calculations with AO basis sets of systematically increasing size (*e.g.* Dunnings[114] correlation-consistent sets (aug-)cc-pVXZ, $X = 2,3,4\cdots$) with a computationally cheap but crude correlation model (typically MP2) and extrapolate the interaction energies to the complete basis set (CBS) limit which yields ΔE_{MP2}^{CBS}. This result is then corrected for higher-order correlation effects and systematic MP2 errors by a CCSD(T) calculation with a small basis set and additivity assumptions

$$\Delta E_{est.CCSD(T)}^{CBS} = \Delta E_{MP2}^{CBS} - (\Delta E_{MP2}^{small} - \Delta E_{CCSD(T)}^{small}). \quad (7)$$

The accuracy of this multi-level approach mainly depends on size of the correction term in parentheses and the quality of the small basis set (which is sometimes chosen too small in order to keep the computation tractable). The estimated accuracy is about 3–5% for ΔE in dispersion-dominated vdW complexes of medium-sized molecules.

With small basis sets a further quite serious complication arises because the description of the monomer wavefunctions is then very incomplete. When forming a complex (*i.e.*, for shorter intermolecular distances), this can variationally be improved by utilizing (unused) basis functions of the interacting partner. This leads to spurious lowering of the interaction energy, referred to as basis set superposition error (BSSE). This error can be approximately removed by using the counterpoise procedure (CP)[115,116] which, however, requires two additional computations and is furthermore not applicable in the intramolecular case. The CP correction is absolutely necessary for correlated calculations on small systems and with small (less than triple-zeta) basis sets. It can be avoided in most DFT calculations with triple-zeta basis sets, for larger systems, when CBS extrapolations are performed, or when local correlation methods[21] are used. If not mentioned otherwise, CP corrections are not performed in this work. We will come back to this point with some examples in Sections 3.2 and 3.3.2, and also refer the reader to the discussion in ref. 31,32.

We exclusively use here AO basis of at least triple-zeta quality, *i.e.*, for the DFT-D calculations mostly those of Ahlrichs (TZV)[117] with two (2d,2p) or three (2df,2pd) sets of polarization functions. Larger sets or those including diffuse basis functions that are important in wavefunction-based methods have been found to be unnecessary in DFT-D calculations[32] when errors for ΔE of about 5–10% are acceptable. We consider standard GGA-type density functionals like BLYP[118,119] or PBE[120] and the new B97-D.[32] The latter is based on a re-parameterization of Becke's ansatz from 1997,[121] but now explicitly by including the $-C_6 \cdot R^{-6}$ dispersion terms. This should by construction avoid double-counting effects of electron correlation, and the density functional description is restricted to short-range electron correlations. In some cases we also present results with the popular B3LYP[122,123] hybrid functional. In the SAPT-DFT computations, we employ the PBE0[124] density functional. When functionals are used together with the dispersion correction, we add the suffix '−D' to the

functional name in order to distinguish from conventional DFT computations (occasionally reported to illustrate the effect of dispersion).

If not mentioned otherwise, the molecular geometries have been fully optimized. For a detailed discussion of the problems with weakly bound complexes in this context and possible sources of numerical error see ref. 31. The interaction energies given are purely electronic, *i.e.*, do not contain vibrational zero-point energies. These corrections from D_e (*i.e.*, $-\Delta E$) to D_0 range from about 0.3 kcal mol^{-1} for the weakly bound benzene dimer to *e.g.* about 2 kcal mol^{-1} for the more strongly bound water dimer.

All DFT-D and spin component scaled (SCS)-MP2 computations were performed with slightly modified versions of the TURBOMOLE suite of programs.[125,126] In all these calculations the resolution of identity (RI) approximation for the two-electron integrals[127,128] is used, which speeds up the computations by a factor of 5–15 with an insignificant loss of accuracy. The RI auxiliary basis sets[129,130] are taken from the TURBOMOLE library.[131] The SAPT computations were performed with MOLPRO[132] and also employ RI (density-fitting) approximations as implemented by Heßelmann, Jansen and Schütz.[133]

3 Results and discussion

3.1 Comparison of DFT-D and SAPT results

In the SAPT method,[39] the interaction energy (up to the second order) is expressed as

$$\Delta E = E_{es}^{(1)} + E_{exr}^{(1)} + E_{ind}^{(2)} + E_{exr-ind}^{(2)} + E_{disp}^{(2)} + E_{exr-disp}^{(2)}. \quad (8)$$

The interaction energy components in eqn (8) are called polarization energies and represent different order corrections in perturbation theory. They have a clear physical interpretation and correspond to the electrostatic (Coulomb interactions of charge densities of the unperturbed monomers), exchange (effect of Pauli repulsion or, equivalently, of anti-symmetrization of the unperturbed wavefunctions of the monomers), induction (interactions of induced multipole moments with permanent moments of the partner, sometimes confusingly called polarization), exchange-induction (effect of anti-symmetrization of induction wavefunctions), dispersion (interaction of instantaneous multipole moments), and exchange-dispersion (effect of anti-symmetrization of dispersion wavefunctions) interactions, respectively. For convenience and to allow comparison with our partitioning approach in DFT-D, $E_{es}^{(1)}$, $E_{ind}^{(2)}$, and $E_{exr-ind}^{(2)}$ are added to yield E_{es}, and $E_{exr-disp}^{(2)}$ and $E_{disp}^{(2)}$ added to yield E_{disp}. The combined discussion of $E_{es}^{(1)}$ and $E_{ind}^{(2)}$ terms as 'electrostatic' is very reasonable for the neutral systems investigated here (as opposed to *e.g.* cation–π complexes) because the induction energy is then typically much smaller compared to $E_{es}^{(1)}$ and furthermore not very system-specific. Note further that in SAPT the complex is never treated as a full system in one computation, and thus this approach has the big advantage of being free of BSSE. On the other hand, only complexes with separable fragments can be treated, which excludes the important intramolecular case. We use SAPT here not as a computational method for practical applications but want to compare the interaction components with those from DFT-D to show that our DFT treatment of weak interactions has a very solid physical basis.

We partition the interaction energy from a supermolecular DFT-D computation by the so-called energy decomposition analysis (EDA) that goes back to the work of Morukuma.[134,135] The EDA has been proven to give detailed information about the nature of chemical bonding,[136] as well as for the interactions in hydrogen-bonded systems[137] and in supramolecular structures.[35] The formation of bonding between two fragments is divided into three physically plausible steps. In the first step, the fragment electronic densities (in the frozen geometry of the super-molecule) are superimposed, which yields the quasi-classical electrostatic interaction energy ($E_{es}^{(1)}$). Renormalization and orthogonalization of the product of monomer wavefunctions yields a repulsive energy term that corresponds to E_{exr}. In the final step, the molecular orbitals are allowed to relax to their final form, which yields the (usually stabilizing) induction energy, and also includes orbital and charge-transfer terms that are in part absent in second-order SAPT. In a manner similar to that above, we add this term to $E_{es}^{(1)}$, which yields then E_{es}. The dispersion energy term (mainly $E_{disp}^{(2)}$ in SAPT) is calculated with the DFT-D approach. The total interaction energy

$$\Delta E = E_{exr} + E_{es} + E_{disp} \quad (9)$$

differs from the true interaction energy only by the energy necessary to bring the optimum monomer geometries into the form they have in the super-molecule. This deformation energy is very small in most cases (<2% of ΔE) and not discussed further. The three terms E_{exr}, E_{es}, E_{disp} and the total interaction energy from EDA and SAPT are compared for three typical (but small) complexes in what follows.

Fig. 5 and 6 show that (although SAPT and EDA are very different approaches) not only the total interaction potentials but also the different parts are quite close to each other. This holds in particular for E_{disp}, which is remarkably similar in both approaches (see Fig. 6). At short distances the effect of the damping function (which is DFT-D-specific to avoid electron correlation double-counting and thus has no analogue in SAPT) is clearly visible. Larger systematic differences between SAPT and EDA are observed for E_{exr} and E_{es}, *i.e.*, the former is higher and the latter is always lower in EDA. The reasons for this are presently not clear and deserve more research. Note, however, that besides an incorrect DFT description of the interaction, other points must also be considered. For example, the AO basis set used in SAPT (aug-cc-pVTZ) is not large enough to provide converged potentials, which is evident from SAPT binding energies that are in general too small. Also, for more polar systems, higher-order effects (*e.g.* charge transfer), which are accounted for in EDA by a full SCF procedure for the complex, are not considered in second-order SAPT. Furthermore, one also has to consider that there is some arbitrariness whether the mixed terms $E_{exr-ind}^{(2)}$ and $E_{exr-disp}^{(2)}$ should be collected in EXR or in ES/dispersion parts, respectively.

In any case, the individual contributions to the interaction potentials easily allow a classification of the main bonding motifs in the three complexes. In the NH_3 dimer, dispersion is qualitatively not important and the potential at larger distances is dominated by E_{es}, which is a signature for a complex of polar molecules with some hydrogen-bonding character. The ethene dimer represents the other extreme where the interaction energy asymptotically is given by E_{disp} while E_{es} is smaller at distances larger than 4 Å. The CH–π complex between ethene and ethyne is intermediate: E_{es} is

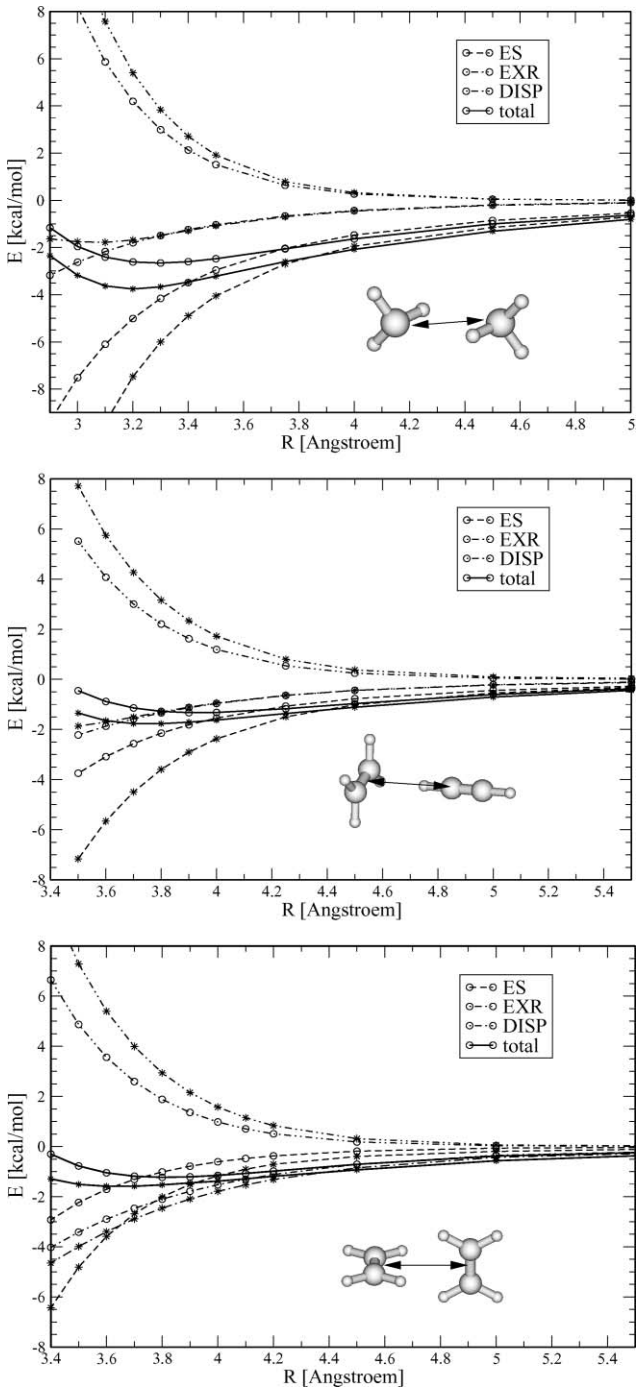


Fig. 5 Comparison of interaction potentials and components from EDA (B97-D/TZV(2df,2pd), marked by stars) and SAPT-DFT(PBE0)/aug-cc-pVTZ (marked by open circles) computations for the ammonia dimer (top), the ethene–ethyne (CH–π, middle) complex, and the ethene dimer (bottom). In all cases fixed (optimized) monomer geometries are used and the intermolecular distance R is defined by an arrow.

very important near the minimum but E_{disp} is still significant, and at larger distances both contribute almost equally to the binding.

3.2 Benchmark sets and model complexes

Before turning to examples of concrete experimental relevance we want to show how accurate the DFT-D method is for typical

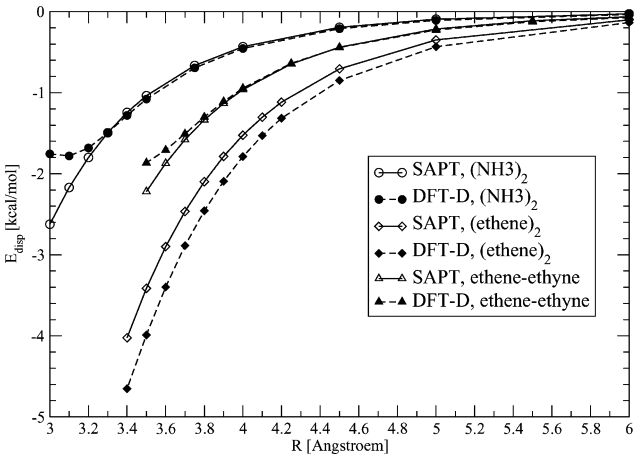


Fig. 6 Comparison of dispersion contributions to the interaction energy from EDA (B97-D/TZV(2df,2pd)) and SAPT-DFT(PBE0)/aug-cc-pVTZ computations for the ammonia dimer, the ethene–ethyne complex, and the ethene dimer.

non-covalent bonding interactions. For this purpose, benchmark sets of molecular complexes are considered. We first investigate a set of systems with biochemical relevance and secondly compare substituted benzene–benzene complexes investigated in detail by Sinnokrot and Sherrill.[11,138]

A database dubbed JSCH-2005 with CCSD(T) complete basis set limit intermolecular interaction energies of 165 non-covalent complexes has been released recently by Jurecka *et al.*[139] The majority of complexes are DNA base pairs (128), but 19 amino acid pairs and 18 other small complexes are included, too. The size of the complexes varies from six atoms in the water dimer to 54 atoms in the phenylalanine–tryptophan pair. Also, the range of interaction energies covered is remarkable: hydrogen-bonded DNA base pairs have an average interaction energy of −21 kcal mol^{-1}, stacked base pairs of −8 kcal mol^{-1}, and inter-strand base pairs of −1 kcal mol^{-1}. For amino acid pairs, the interaction energies reach values up to −113 kcal mol^{-1} between oppositely charged glutamic acid and lysine residues. This set has recently been investigated successfully with the improved DFT-D method and several common functionals.[37] We present here only data for a smaller subset of complexes that nevertheless covers most types of non-covalent interactions in bioorganic chemistry. The results for interaction energies with B97-D and BLYP-D methods are shown in Table 1. Note that the reference values also have an estimated error of 2–3% of ΔE. A close inspection of the data in Table 1 reveals that both DFT methods are very accurate. The root-mean-square deviations of the DFT-D ΔE values from the reference are 0.6 and 0.5 kcal mol^{-1} for the BLYP-D and B97-D functionals, respectively. For the typical binding energies of −10 to −20 kcal mol^{-1}, this corresponds to less than 5% of ΔE. The differences between the largest positive and the largest negative deviations are 2.0 and 2.1 kcal mol^{-1}. For the B97-D functional, the largest positive deviation occurs for the hydrogen-bonded uracil dimer with 1.2 kcal mol^{-1}, which is only 6% of the reference value. The largest negative deviation is obtained for the benzene–H_2O complex, and amounts to −0.86 kcal mol^{-1}, which is the largest relative deviation (26%) in this small dataset found with B97-D. Similarly small errors (root-mean-square deviations of about 1 kcal mol^{-1}) have also been observed for the full data

Table 1 DFT-D/TZV(2df,2pd) intermolecular interaction energies and deviations with respect to reference values (in kcal mol^{-1})

No.	Complex (symmetry)	B97-D ΔE [a]	B97-D CP [b]	BLYP-D ΔE [a]	Est. CCSD(T)/CBS ΔE_{ref} [c]
Hydrogen-bonded complexes					
1	$(NH_3)_2$ (C_{2h})	−3.72	−0.44	−4.16	−3.17
2	$(H_2O)_2$ (C_s)	−5.07	−0.68	−5.80	−5.02
3	Formic acid dimer (C_{2h})	−18.25	−0.64	−19.34	−18.61
4	Formamide dimer (C_{2h})	−15.28	−0.58	−16.39	−15.96
5	Uracil dimer (C_{2h})	−19.45	−0.45	−20.73	−20.65
6	2-Pyridoxine–2-aminopyridine	−17.13	−0.59	−18.05	−16.71
7	Adenine–thymine WC	−16.20	−0.58	−17.19	−16.37
Complexes with predominant dispersion contribution					
8	$(CH_4)_2$ (D_{3d})	−0.57	−0.02	−0.36	−0.53
9	$(C_2H_4)_2$ (D_{2d})	−1.55	−0.05	−1.55	−1.51
10	Benzene–CH_4 (C_3)	−1.51	−0.04	−1.37	−1.50
11	Benzene dimer (C_{2h})	−2.67	−0.22	−2.35	−2.73
12	Pyrazine dimer (C_s)	−4.07	−0.25	−4.05	−4.42
13	Uracil dimer (C_2)	−10.02	−0.63	−10.50	−10.12
14	Indole–benzene	−4.72	−0.38	−4.55	−5.22
15	Adenine–thymine stack	−12.11	−0.80	−12.85	−12.23
Mixed complexes					
16	Ethene–ethyne (C_{2v})	−1.73	−0.07	−1.62	−1.53
17	Benzene–H_2O (C_s)	−4.14	−0.71	−4.16	−3.28
18	Benzene–NH_3 (C_s)	−2.75	−0.35	−2.66	−2.35
19	Benzene–HCN (C_s)	−4.88	−0.08	−4.87	−4.46
20	Benzene dimer (C_{2v})	−2.93	−0.13	−2.76	−2.74
21	Indole–benzene T-shape	−6.26	−0.30	−6.16	−5.73
22	Phenol dimer	−6.60	−0.55	−7.35	−7.05

[a] Not corrected for BSSE. [b] Counterpoise correction. [c] Estimated CCSD(T)/CBS.[139]

set of 165 complexes.[37] Noteworthy is also the very consistent description of complexes of different bonding type, *e.g.* H-bonded compared to π-stacked structures. The second column of the table also includes the counterpoise correction for BSSE, which is in all cases very small (absolute value <1 kcal mol^{-1}) and negligible for the larger molecules in particular. Typically, the CP correction with properly polarized triple-zeta basis sets is <5% of ΔE, which is on the order of normal basis set effects and which does not warrant the additional CP computations.

Our second test consists of hetero-dimers of benzene with substituted benzenes (R = CH_3, OH, F, CN) in stacked (face-to-face) and T-shaped conformations (T with the substituted benzene as the CH donor and T2 in reverse mode). These systems have been used to understand the substituent effects on π–π interactions[11] (for the analysis of a similar pyridine-substituted benzene model system see ref. 140,141). The theoretical results of Sinnokrot and Sherrill (est. CCSD(T)/aug-cc-pVTZ), which contradict the empirical rules of Hunter and Sanders,[142] are fully supported by the present B97-D calculations. Independent of the nature of the substituent (electron-donating or -withdrawing) all stacked dimers are more strongly bound than the unsubstituted benzene dimer. For the T-shaped arrangements, CH_3/OH and F/CN substituents fall in two classes, as expected when ES effects are dominant (see below).

As it is evident from Fig. 7, the DFT-D method provides very accurate binding energies, and not only on an absolute scale; the substituent effects are also described very well. The systematically stronger binding provided by B97-D (*i.e.*, all data points are located below the dashed line by about 0.3 kcal mol^{-1} on average) can be traced back to under-binding of the reference method that used only an aug-cc-pVTZ AO basis. We also decomposed the binding energies using EDA, and plot in Fig. 8 the energetic changes compared to the corresponding benzene dimer due to EXR, ES, and dispersion contributions.

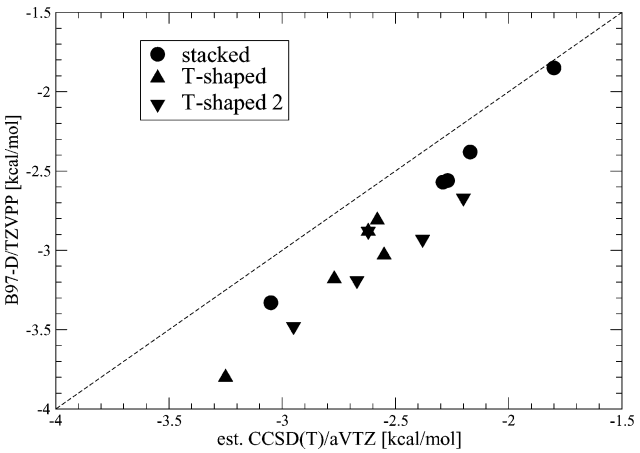


Fig. 7 Comparison of binding energies (frozen monomer geometries) for benzene-substituted benzene (R = H, OH, CH_3, F, CN) from B97-D/TZV(2df,2pd) and estimated CCSD(T)/aug-cc-pVTZ methods. The dashed line has a slope of unity and an intercept of zero.

This analysis, that is also relevant to intramolecular cases as in the triptycene derivatives discussed in Section 3.3.1, reveals the different nature of binding in stacked and T-shaped structures. In agreement with the results of Sinnokrot and Sherrill, we see for the stacked forms that the ES contribution is more stabilizing for all substituents. The very strong binding for R = CN can be explained by better ES interactions compared to the other substituents, while

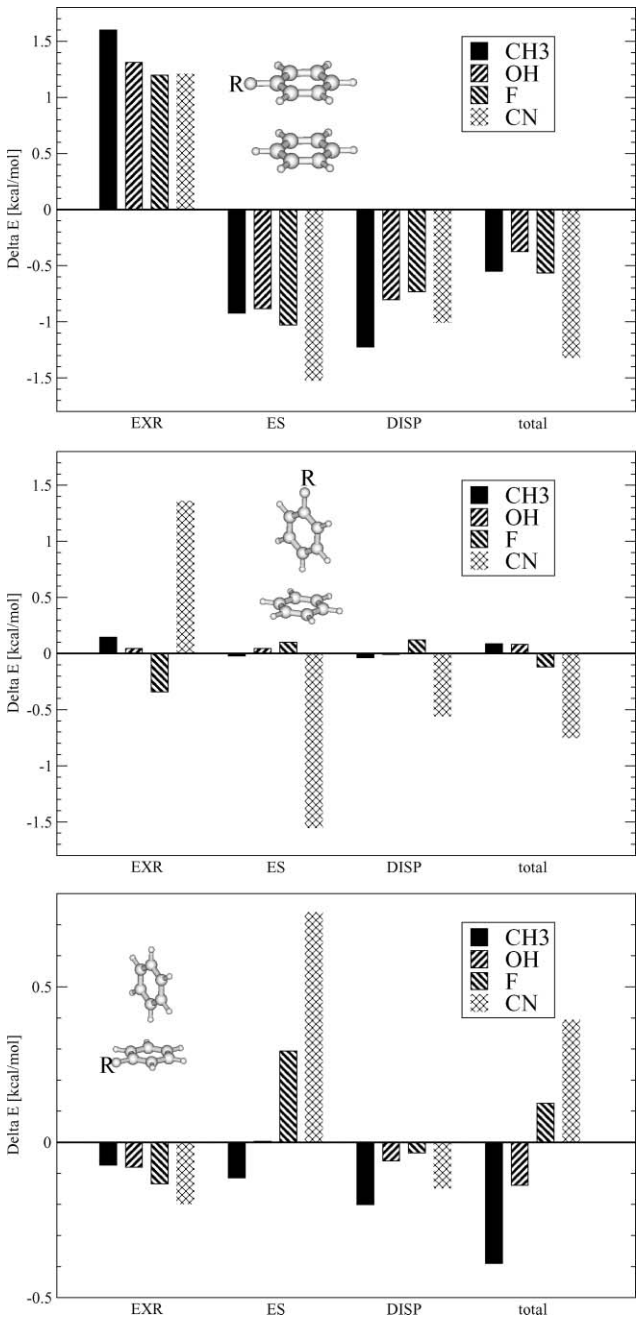


Fig. 8 Contributions to the interaction energy relative to the benzene dimer for heterodimers with one mono-substituted benzene (R = OH, CH_3, F, CN) in the stacked (top), T-shaped (middle) and T-shaped(2) (bottom) configuration. The components to ΔE are computed by EDA at the B97-D/TZV(2df,2pd) level using geometries from ref. 11.

the relatively good binding for R = CH_3 mainly results from a large dispersion term. For both T-shaped arrangements (and for T(2) in particular) one can see the dominant contribution of ES for the substituent effect, which emphasizes the simple picture of the interaction of the slightly polar CH donor bond with the π-density of the acceptor ring. The exceptionally large difference of ES and EXR components in the T-shaped case for R = CN is mainly attributed to the short inter-ring distance of 4.9 Å instead of 5.0 Å as for the other complexes.

Before discussing the results for intramolecular examples, we want to comment on the commonly used terms 'π–π stacking' or 'π–π interactions', which are rather mysterious and are often used when a deeper understanding of the system is missing (see also the comment in ref. 11). In our opinion, the term π–π stacking should merely be used as a geometrical descriptor when aromatic or other unsaturated organic molecules have their molecular planes in a more-or-less parallel orientation. In essence, our view is that special π–π interactions in aromatic vdW complexes simply do not exist (for borderline cases see ref. 143). Let us note first that *e.g.* the stacked (parallel-displaced) benzene dimer has an even smaller binding energy (about −2.8 kcal mol^{-1}) than the completely saturated pentane dimer (about −3.9 kcal mol^{-1}),[144] which has the same size. This result is incompatible with the assumption of special π–π interactions. Instead, both systems (unsaturated and saturated) are typical vdW complexes where dispersion is absolutely essential for the binding. For larger aromatic systems, the stacked orientation just minimizes the interatomic distances for optimal dispersion interactions. Because in this arrangement not many orbitals overlap (just significantly the π-MOs), the EXR is quite small (*e.g.* compared to the pentane dimer or the T-shaped arrangement), and this helps to compensate for unfavorable ES effects in this orientation.

The most special property of π-systems compared to saturated systems is their better interaction with polar molecules, as illustrated by the comparison of the electrostatic potentials (ESP) of benzene and pentane in Fig. 9.

It is clearly seen that benzene has pronounced negative regions above and below the molecular plane that are attributed to the π-electrons. These regions may serve *e.g.* as an H-bond acceptor similar to the lone-pairs of oxygen and nitrogen, which explains the quite strong binding of benzene with water and ammonia (see Table 1 and also Section 3.4.1).

Such regions are obviously missing in pentane, which only has very shallow minima in the ESP of about −3 kcal mol^{-1} depth. These plots also help to understand why an EDA of the interaction energy terms reveals the electrostatics to be more attractive in the pentane dimer than in the benzene dimer. They further explain why the arene units in vdW complexes are mostly displaced with respect to each other, which minimizes the repulsive ES interactions between the mid-parts of the rings.

3.3 Examples for intramolecular interactions

3.3.1 'π-Stacking' in triptycene derivatives. The equilibrium between *syn*- and *anti*-forms of the triptycene derivative shown in Fig. 10 have been investigated by Gung *et al.*[145] using NMR spectroscopy in chloroform. The two conformers differ mainly by the orientation of the $–CH_2CPh$ ring with respect to the R–PhCOO moiety. The two aromatic units in the *syn*-form are located in a stacked, parallel-displaced arrangement. This inter-ring interaction, which also includes components from the carbonyl moiety should, however, stabilize the *syn*-conformer relative to the *anti*-conformer; a result which is indeed found experimentally.

The experimental ΔG values for the *anti* → *syn* rotation have been obtained for different *para*-substituents R on the PhCOO moiety (a further substituent on the experimentally investigated compounds in *para*-position to the PhCOO group has been discarded for the calculations). The experimental values of the

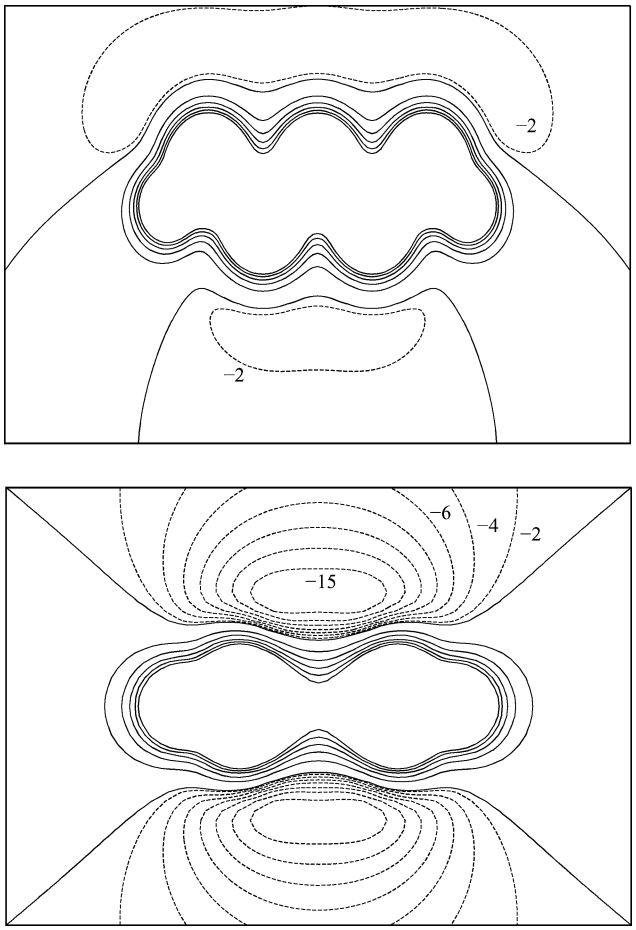


Fig. 9 Contour-line plot of the electrostatic potential of pentane (top, in the CCCCC plane) and benzene (bottom, orthogonal to the molecular plane). Dashed lines indicate a negative (positive-charge-attracting) potential in steps of ±2 kcal mol^{-1}.

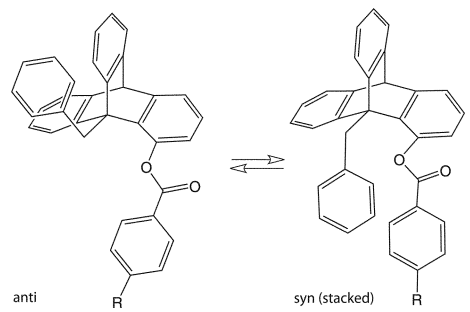


Fig. 10 Structures of the two conformers of the triptycene derivatives.

compounds considered here (R = NO_2, CN, F, H, Me) range from −0.69 kcal mol^{-1} (stacked form more stable) for R = NO_2 to 0.14 kcal mol^{-1} (R = Me). It is important to note that these values have been obtained in chloroform solution and thus contain solvation effects that are not fully considered in the our gas-phase dielectric continuum model (COSMO[146]) calculations. While in the *anti*-form both arene rings are expected to have an almost complete solvation shell, this is in part lost in the stacked arrangement. Thus, compared to the gas phase, the *anti*-form is stabilized in solution more than the *syn*-form. Therefore, our theoretical results

Table 2 Comparison of theoretical[a] and experimental conformational energies *anti* → *syn* (kcal mol^{-1}). Values in parentheses are those without the dispersion correction

R	Δ*E*	Δ*E*(shifted) [b]	Δ*G*(exp) [c]
NO_2	−2.02 (2.44)	−1.03 (−0.37)	−0.69
CN	−1.80 (1.99)	−0.81 (−0.82)	−0.55
F	−1.24 (2.81)	−0.25 (0.00)	−0.14
H	−0.97 (2.83)	0.02 (0.02)	0.02
Me	−0.96 (2.55)	0.03 (−0.26)	0.14

[a] B97-D + COSMO(ε = 4.81)/TZV(2d,2p) level. For all atoms of the triptycene moiety, a smaller SV(d,p) AO basis set is used. [b] 0.99 kcal mol^{-1} or −2.81 kcal mol^{-1} (pure DFT) added (values derived from R = H) to account for the solvent effect. [c] In $CDCl_3$, error ±0.05 kcal mol^{-1} (ref. 145).

are shifted with respect to the experimental value by the difference in the free solvation enthalpies. This shift is estimated from the difference between Δ*E*(calc) and Δ*G*(exp) for R = H, and these corrected values are shown in the second column of Table 2. Perusing Table 2, one finds an excellent agreement between the corrected theoretical values and experiment, *i.e.*, the substituent effect is described accurately to within ±0.3 kcal mol^{-1}. The solvent effect of about 1 kcal mol^{-1} seems reasonable for the loss of one solvation shell for a substituted benzene in chloroform. Note that the DFT values without the dispersion correction are much too high (the *anti*-form is too stable by 3–4 kcal mol^{-1}), and furthermore the computed substituent effects are incorrect for R = NO_2 and R = Me. This underlines that only a comprehensive treatment of all interaction terms including dispersion can provide quantitatively correct non-covalent interactions. Note also that much of the success for this system results from the complete DFT-D geometry optimizations that could be performed. Without the dispersion correction, very distorted geometries are obtained that are useless for the evaluation of the conformational energies. This is an appealing feature of the DFT-D method; the dispersion term simply can be switched off and the effects on geometry or energy can be monitored. This is shown for example in Fig. 11 for the compound with R = H.

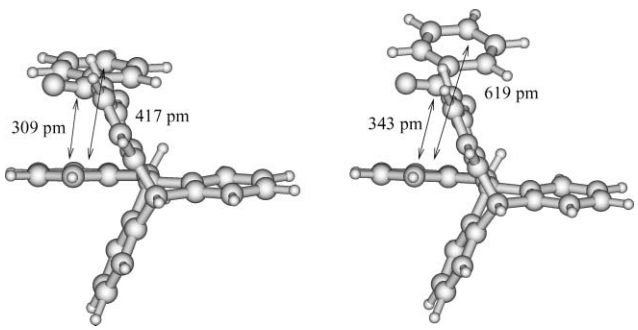


Fig. 11 Side view of the optimized structures (B97-D/TZV(2d,2p)-SV(p,d) of the *syn*-conformer (R = H) with (left) and without (right) the dispersion correction. The arrows indicate the distance between the *para*-carbons of the two rings (417 and 619 pm) and the carbonyl carbon atom and the *ortho*-carbon of the lower ring (309 and 343 pm).

The inter-ring distances that are relevant in the stacking depend dramatically on the dispersion contribution, as can be seen by lengthening of the carbonyl-group–ring distance by 34 pm and between the two *para*-carbon atoms of more that 200 pm. Without

the dispersion potential, the strongly repulsive EXR terms are not compensated for and the entire R–PhCOO moiety tries to avoid the ring by bending up and undergoing torsion. This example shows clearly that with conventional functionals qualitatively wrong geometries can be obtained.

3.3.2 Folding in alkane chains. The folding of molecular structures from spatially extended linear arrangements to those that are more dense and space-filling is an important elementary process in proteins (for a related study on α- *vs.* 3_{10}-helix conversion see ref. 147). We study the energetic consequences of a simple folding event here in a non-polar model system in order to monitor in particular the effects of dispersion.

For small chain lengths, non-branched alkanes C_nH_{2n+2} are most stable in the linear (L), C_{2h}-symmetric form (*n* even). In the case of butane, the relative energy to the higher-lying *gauche* minimum is about 0.9 kcal mol^{-1}.[148] When the alkane becomes longer, a *gauche*-type torsion in the middle of the chain can lead to an arrangement of two spatially close-lying 'arms' of the molecule, as shown for example for $C_{30}H_{62}$ in Fig. 12.

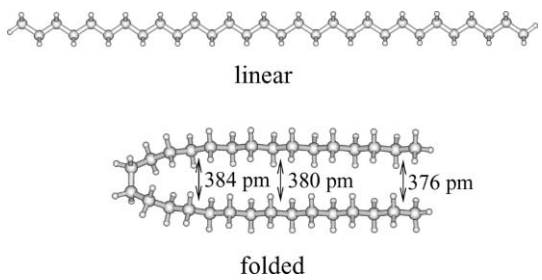


Fig. 12 Optimized structures (BLYP-D/TZV(d,p)) of the linear (C_{2h}) and folded (C_2) conformers of $C_{30}H_{62}$.

The intramolecular interactions between these parts can over-compensate the linear–*gauche* energy difference for larger chains, so that folded alkanes become more stable than linear conformers. We investigate here the energy difference between linear (L) and *gauche* (singly-folded, F) forms ΔE_{LF} as a function of the chain length (n = 14, 22 and 30). Full geometry optimizations were performed at the BLYP-D/TZV(d,p) level. Single-point energies were obtained with several functionals, and the results are compared to those from the MM3 force field[149] and the MP2 method (Table 3). Because the counterpoise correction (important in MP2 calculations) cannot be applied here, we present MP2 results with two AO basis sets of different size. Note that the calculation for $C_{30}H_{62}$ with the large aug-cc-pVTZ AO basis involves about 3000 orbitals, which is not far from the limit of what routinely can be carried out on today's computer hardware.

Having already pointed out in the previous examples the importance of dispersion for larger molecules, it comes as no surprise that the effect for these long alkane chains is dramatic. With HF or BLYP, which neglect vdW interactions completely, the linear form becomes increasingly more stable as the chain length grows. For $C_{30}H_{62}$, the folded conformer is between 20 and 30 kcal mol^{-1} less stable, which can be mainly attributed to the intramolecular EXR between the chains. This is drastically different with methods that account for dispersion interactions, where even for n = 14 the linear form is slightly more stable. The crossing point is between n = 14 and n = 22 where the folded forms become more stable. For the longest chain, the energy gain due to folding is very significant and amounts to 5–10 kcal mol^{-1}. Note also that for the smallest system, HF and BLYP relative energies deviate by 5–7 kcal mol^{-1} from the other methods. We also tested a common force field (MM3)[149] for that problem and found a reasonable but not perfect agreement with the quantum chemical data. The basis set effect when going from the TZV(2df,2pd) to the aug-cc-pVTZ AO basis is in the expected direction (stabilizing the folded forms due to a better description of intramolecular dispersion) but the changes are relatively small. Tentatively, this can be explained by a compensating effect of larger BSSE and missing diffuse functions (to describe dispersion and induction) in the TZV(2df,2pd) calculation. Note also the almost perfect agreement between the B97-D and MP2/aug-cc-pVTZ values.

Table 3 Energy difference[a] between linear and folded forms ΔE_{LF} (in kcal mol^{-1})

Method	ΔE_{LF}		
	$C_{14}H_{30}$	$C_{22}H_{46}$	$C_{30}H_{62}$
HF [b]	−8.9	−23.8	−30.6
BLYP [b]	−6.9	−16.5	−20.8
BLYP-D [b]	−1.4	6.8	12.7
B97-D [b]	−1.5	4.8	9.8
MP2 [c]	−2.6	1.8	5.9
MP2 [d]	−2.2	3.6	8.8
Force field (MM3) [e]	−5.1	−2.9	12.7

[a] BLYP-D/TZV(d,p)-optimized structures. A negative sign indicates that the linear form is more stable. [b] TZV(2d,2p) AO basis. [c] TZV(2df,2pd) AO basis. [d] aug-cc-pVTZ AO basis. [e] Fully optimized.

3.3.3 The dimerization of anthracene. The photo-dimerization of anthracene, a clean and reversible reaction that yields in a [4 + 4] cycloaddition manner the covalently bound polycyclic dimer, has been known for about 140 years[150] (see Fig. 13). The dimer is thermally labile at elevated temperatures and reverts back to the two monomers, which has raised hopes for its use in solar energy storage devices.

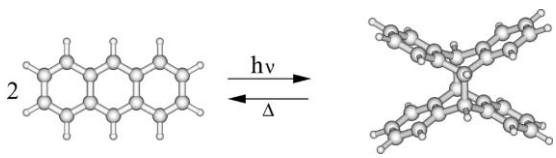


Fig. 13 Dimerization and thermal dissociation of anthracene.

The dissociation energy (D_e) of the dimer has recently been investigated in great detail theoretically in ref. 151, in which experimental solution data have also been discussed. Here we just want to compare DFT-D data with the most accurate theoretical reference value of about 9 kcal mol^{-1} for D_e that has been obtained with coupled-cluster and quantum-Monte-Carlo-based methods.[151]

The results of calculated D_e values are summarized in Table 4. Quite surprisingly for a seemingly simple organic reaction, all density functionals except PBE-D yield the wrong sign for D_e, *i.e.*, two anthracene molecules are computed to be more stable than the dimer. Particularly striking is the bad performance of the commonly used BLYP and B3LYP functionals, which lead to huge and unacceptable errors of 35–45 kcal mol^{-1}. The first hint to

Table 4 Dissociation energies D_e (in kcal mol^{-1}) for the anthracene dimer

Method[a]	AO basis	D_e
B3LYP	cc-pVTZ	−25.8
BLYP	TZV(2df,2pd)	−35.3
HF	cc-pVTZ	−20.7
PBE	TZV(2df,2pd)	−14.5
BLYP-D	TZV(2df,2pd)	−9.1
B97-D	TZV(2df,2pd)	−7.2
PBE-D	TZV(2df,2pd)	1.9
SCS-MP2	cc-pVTZ	13.4
MP2	cc-pVTZ	21.4
Best theoretical estimate (ref. 151)		9 ± 3

[a] Single point calculations using the MP2/TZV(2d,2p) geometry.

the problem of the discrepancy between wavefunction theory (*e.g.* SCS-MP2[152]) and DFT comes from the large dispersion correction obtained from the DFT-D method. Also, the large D_e value obtained by MP2 (which overestimates dispersion interactions involving π-systems; see *e.g.* ref. 153,154) is in line with the interpretation that dispersion is at the heart of the problem. This can easily be understood by considering the large size of the molecules (especially the dimer), with many interatomic distances close to those of typical vdW minima. When used together with the PBE density functional, the new dispersion correction lowers the dimer relative to two separated anthracene molecules by about 16 kcal mol^{-1}, leading to a D_e value with a correct sign. Obviously (and ultimately quite understandably) there are strong intra-molecular vdW interactions in the dimer. From the structure, one can clearly see that the four outer benzene rings are in a stacked arrangement (C–C distances are between 270 and 450 pm), which should add roughly 5–10 kcal mol^{-1} internal stabilization compared to two monomers. These effects are entirely due to long-range electron correlation, and thus are absent in HF and standard DFT, which both yield a dimer that is too unstable ($D_e < 0$), while all correlated wavefunction methods that accurately include them yield $D_e > 0$. Note that the wrong sign provided by B97-D and BLYP-D is mainly attributed to an incorrect description of the σ–π transformation and the aromatization energies in the reaction. It is clear that the DFT-D model can only improve the long-range part of the dispersion interactions, and in such a case medium range-correlation effects also play a major role. In any case, this example shows clearly how important intra-molecular dispersion effects in reactions with medium-sized molecules really are, and furthermore, the limits of the simple (orbital-independent) DFT-D treatment.

3.4 Examples for intermolecular interactions

3.4.1 Complexes of polar aromatic molecules: corannulene. Van der Waals complexes of polar aromatic molecules are particularly interesting because of a subtle balance between dispersion and electrostatic effects. This has previously been investigated for pyridine[34] and azulene[36] dimers. One of the most interesting and exotic hydrocarbons in this context is corannulene. Because of the (geometrical) ring-strain it is non-planar and has a bowl-shape form. Furthermore, due to the 20 π-electrons it does not obey the Hückel $4n + 2$-rule. Nevertheless, it benefits from a reasonable resonance energy stabilization which can be explained by, for example, the mesomeric structures shown in Fig. 14.

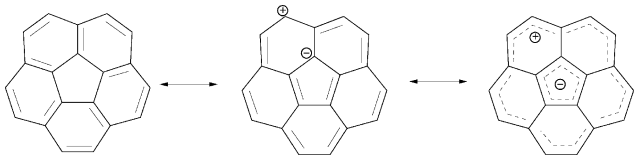


Fig. 14 Resonance structures of corannulene.

Another important characteristic of this molecule is the high dipole moment of about 2.1 debye (B97-D/TZV(2d,2p) level) caused by its vaulted shape and the electron distribution that reflects the significant participation of the aforementioned mesomeric structures. This makes corannulene a candidate for relatively stable vdW complexes, as considered in an early theoretical study of its dimer by Tsuzuki *et al.*[155]

Because this molecule can serve as a model for fullerenes and nanotubes, we initiated a systematic study of corannulene complexes with a series of small molecules (H_2O, NH_3, and CH_4). Corannulene has two distinct binding sites at the inside (i) and at the outside (o) of the bowl. As can be seen in Fig. 15, the electrostatic potential of corannulene is more negative on the outside (minimum value of about −12 kcal mol^{-1}) than on the inside (minimum value of about −8 kcal mol^{-1}), which should be compared to the corresponding value of about −10 kcal mol^{-1} for the similarly sized (but planar) coronene. Tentatively, this can explain the slight preference for outside-complexes of corannulene with (electron deficient) transition metal fragments observed in recent X-ray studies.[156] Fig. 16 shows the optimized geometries of

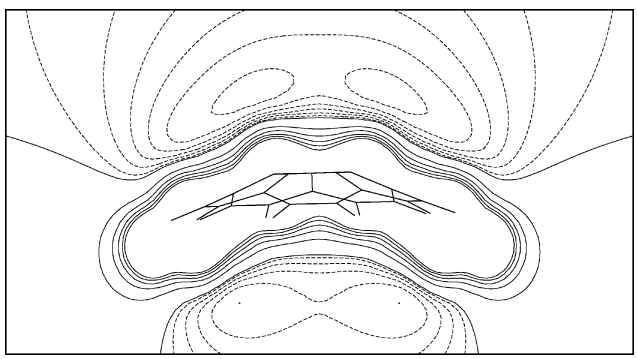


Fig. 15 Contour-line plot of the electrostatic potential of corannulene on a plane containing the C_5 axis through the middle of the molecule. Dashed lines indicate a negative (positive-charge-attracting) potential. The step between the contour-lines is −2 kcal mol^{-1}. The global ESP minimum on the outer side of the bowl is −12.4 kcal mol^{-1}.

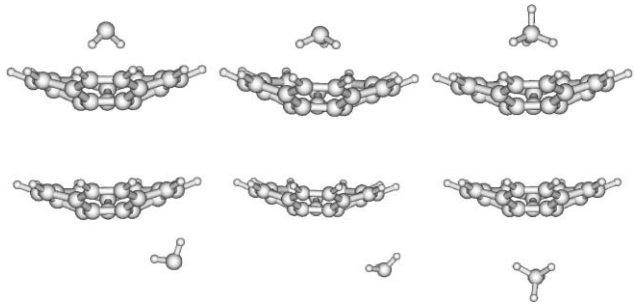


Fig. 16 Geometries of inside-bound (top) and outside-bound (bottom) corannulene complexes with (left to right) water, ammonia, and methane.

Table 5 Energy contributions to binding from EDA (BLYP-D/TZV(2d,2p)) for the different corannulene complexes and total binding energies from DFT-D and SCS-MP2 (all data in kcal mol^{-1}). The values in parentheses are without the dispersion correction

	DFT-D				SCS-MP2[b]
Complex with: [a]	E_{exr}	E_{es}	E_{disp}	ΔE [c]	ΔE
CH_4 (i)	18.0	−12.5	−11.0	−5.5 (5.5)	−4.3
NH_3 (i)	16.4	−13.1	−10.5	−7.3 (3.3)	−4.8
H_2O (i)	11.5	−10.6	−8.6	−7.7 (0.9)	−4.6
CH_4 (o)	6.0	−3.9	−3.9	−1.8 (2.1)	−1.6
NH_3 (o)	6.9	−6.3	−3.8	−3.2 (0.6)	—
H_2O (o)	8.8	−9.5	−4.3	−5.0 (−0.7)	—

[a] Abbreviations: (i): inside; (o) outside. [b] CP-corrected aug-cc-pVTZ values using BLYP-D/TZV(2d,2p) optimized geometries. [c] Fragment deformation energies not included.

all six complexes considered. In Table 5 the corresponding binding energies are collected, and also includes SCS-MP2 values for comparison. This method yields the same trends and qualitative picture, but the inner values are 2–3 kcal mol^{-1} smaller than with DFT-D. A slight overbinding of water and ammonia by DFT-D is noted even for benzene, which may be attributed to a charge-transfer character that is slightly too large; this is a general DFT problem. The SCS-MP2 values on the other hand may be slightly too small due to basis set incompleteness effects.

The structures of the complexes are geometrically quite similar to each other. The hydrogen atoms always point towards the rings, and on the outside the small molecules tend to choose the 'wall' rather than the 'bottom'. Concerning the special shape of corannulene, all three molecules prefer the inside of the bowl. The binding energies are 2–4 kcal mol^{-1} lower than on the outside. This is somewhat counter-intuitive because one would expect the outer rim to be a better hydrogen bond donor (*cf.* Fig. 15). Note also the intriguing result that NH_3 and H_2O bind almost equally strongly (but only inside), while for *e.g.* benzene the binding linearly increases from CH_4 to NH_3 to H_2O.

For further understanding an EDA has been performed, which is also given in Table 5. Obviously, all inside complexes are stabilized more by dispersion than the outside ones, *i.e.*, at the pure DFT level the outside is always energetically favored and the ratio E_{es}/E_{disp} is also larger outside (at least for H_2O and NH_3). We note in passing that without the dispersion correction only H_2O is bound (at the outside). In agreement with the conclusions drawn from the ESP plot, the ES contributions on the outside increase from CH_4 to H_2O, while they are almost constant for the inside structures (the larger absolute values are due to the shorter intermolecular distances). The unusual result that ammonia forms about the same (or an even stronger) bonds than water can again be attributed to the large contribution of dispersion that is: i) more important for NH_3 and CH_4 than for H_2O (decreasing polarizability); ii) that is clearly stronger inside with more shorter interatomic contacts; and iii) that overcompensates the relative ES effects.

3.4.2 Dimerization of large π-systems. Large organic π-systems have a tendency to form aggregates (dimers) even in solution under ambient conditions. The corresponding spectroscopic consequences in the case of merocyanine dyes have been investigated in detail by Würthner *et al.*[157] We take here one of his examples, to show how difficult structure calculations with other methods are. The formula of the investigated dye is shown in Fig. 17 together with results from DFT-D and AM1[158] optimizations of its dimer.

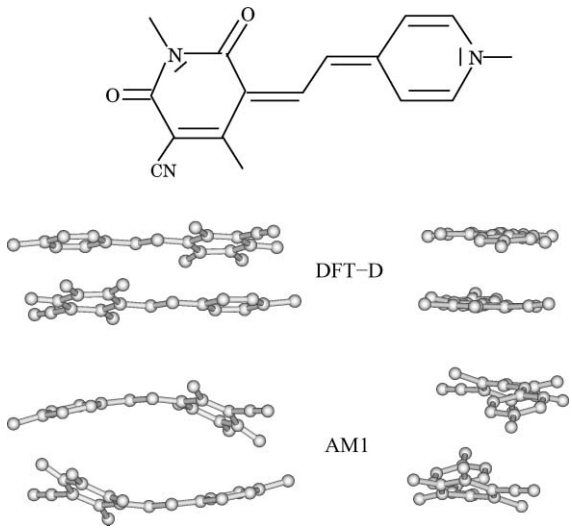


Fig. 17 Formula of the investigated merocyanine dye and two views of the optimized structures (B97-D/TZV(d,p), top) and AM1 (bottom).

The monomer has a large dipole-moment of 18.6 debye (B97-D/TZV(2d,2p)) and thus, the preferred arrangement of the molecules in the dimer is the anti-parallel orientation with C_i symmetry. The monomers are not fully face-to-face but slightly displaced relative to each other, as found in many other aromatic dimer complexes.[159] Their relatively strong interaction is indicated by inter-plane distances between 3.2 and 3.3 Å and a small distortion of the inherently planar chromophore. These theoretical results are in qualitative agreement with the spectroscopic measurements[157] as well as (quite crude) MP2/6-31G(d) optimizations.[160]

Such large π-systems represent difficulties for simpler computational methods. Force-fields are in general not applicable because they can not account for the special unsaturated (delocalized) character that leads to a very non-uniform charge distribution. The only alternatives are simpler MO methods like AM1[158] but these fail completely in this case, as it is evident from the very distorted geometries shown in the bottom of Fig. 17. With AM1 the monomers are shifted with respect to each other, there is a strong bend along their long axis in order to avoid contact that is too close and obviously, the system is held together mainly by Coloumbic forces from the ends of the units.

These failures are understandable by considering the results from an EDA at the DFT-D level. Due to the polar character of the merocyanine dye, one intuitively expects ES to be very dominant for the binding. This picture is more-or-less wrong, *i.e.*, the large E_{es} term of −68.8 kcal mol^{-1} is almost completely quenched by EXR (65.9 kcal mol^{-1}), such that only a very weak interaction of −2.9 kcal mol^{-1} remains at the pure DFT level. The major part of the very large total interaction energy of −43.8 kcal mol^{-1} (−40.5 kcal mol^{-1} including the fragment deformation) is due to dispersion, which stabilizes the dimer relative to the monomers by about 41 kcal mol^{-1}. This is absent in AM1, leading to a

very unbalanced treatment of the interaction terms and finally to completely unreliable structure predictions. Note also that such complex π-systems with many heteroatoms lead to complicated ES interactions that may not be modelled correctly by semiempirical approximations. In any case, the huge dimerization energy of −41 kcal mol^{-1} (about half the dissociation energy of a covalent C–C bond) is a good example how misleading the term 'weak interaction' is, even for medium-sized systems.

3.4.3 Adsorption of methane on carbon nanotubes. The adsorption of small molecules on graphene sheets and carbon nanotubes makes these carbon allotropes attractive as potential gas storage materials. The best estimate for the adsorption enthalpy of methane on graphene under low pressure (low coverage) conditions is −3.0 kcal mol^{-1}.[161] The binding on carbon nanotube material was found to be 76% larger (−5.1 kcal mol^{-1}).[162] The better binding on the tubes was explained by adsorption in channels of nanotube bundles and interaction with more than one (outer) cylinder surface.

We have performed DFT-D calculations (B97-D/TZV(d,p)) to estimate the absolute and relative adsorption energy of methane on a planar polycyclic aromatic compound (coronene) and on the surface and in the inner void of a small single-walled carbon nanotube. Our model tube has a length and diameter of about 15 and 8.4 Å respectively, and comprises 144 carbon atoms. The adsorption energy of one molecule of CH_4, centered on coronene, amounts to −3.1 kcal mol^{-1}, a value that agrees very well with that from experiment.

The optimization of the nanotube complexes (see Fig. 18) revealed adsorption energies of −2.2 kcal mol^{-1} (ΔE_{ads}^{out}) when CH_4 is attached to the outer surface of the tube, and −10.1 kcal mol^{-1} (ΔE_{ads}^{in}) when CH_4 is located inside the tube. The outside binding energy is smaller than that for coronene. Interestingly, for the outside complex, the contribution of ΔE_{disp} to ΔE_{ads}^{out} is only −3.8 kcal mol^{-1}, which is less than in the coronene complex (−5.8 kcal mol^{-1}). For the inside complex, the contribution of ΔE_{disp} to binding is −12.7 kcal mol^{-1}, which explains the large total interaction energy. Thus, the electrostatic/induction terms induced by the curved surface seem to be similar inside and outside, which is corroborated by an ESP plot (not shown). Our calculations for ΔE_{ads}^{out} are in good agreement with experimental reference enthalpies. That our value is smaller supports the conclusion from experiment that more than one tube surface is (on average) involved. Note that without the dispersion correction, the methane molecule is in any case unbound, and furthermore, the large difference between inside and outside situations disappears.

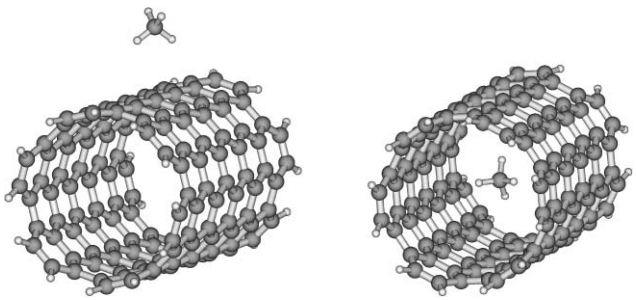


Fig. 18 Model complex of methane with a [6,6] carbon nanotube (B97-D/TZV(d,p)). Left: CH_4 adsorbed on the outer surface, ΔE_{ads}^{out} = −2.2 kcal mol^{-1}. Right: CH_4 located inside the tube, ΔE_{ads}^{in} = −10.1 kcal mol^{-1}.

3.4.4 Gas hydrates. Clathrate hydrates are inclusion compounds consisting of guest molecules (such as noble gases or hydrocarbons) in an ordered network of water molecules. Methane hydrate is of special interest as it may represent a large reservoir of fossil fuel that could be exploited in the future.

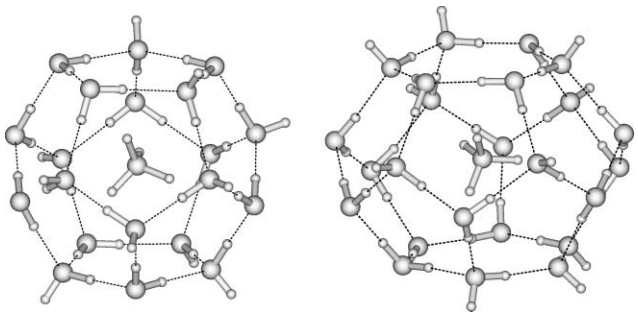


Fig. 19 Model complexes of methane in structure I gas hydrate. Left: CH_4 in a 5^{12} cage ($(H_2O)_{20}$). Right: CH_4 in a $5^{12}6^2$ cage ($(H_2O)_{24}$).

The most abundant structure of methane hydrate (type I) has been found to contain a pentagonal dodecahedron (5^{12} cage) and a tetrakaidecahedron ($5^{12}6^2$ cage).[163] These two cages are used here as methane hosts in supermolecular DFT-D computations (Fig. 19). As the two cages have many isomers with various permutations of internal and external hydrogen atoms, we have proceeded by optimizing only one arbitrarily chosen isomer of each cage with an included methane molecule. Subsequently, single point calculations and structure optimizations of the empty water cage were performed. We compare the energy of the optimized, empty water cage with the energy of the methane complex and the energy of the water cage after CH_4 removal (see Table 6). Here, ΔE is the binding energy of the guest (compared to the optimized empty water cage) and ΔE_{disp} the corresponding dispersion contribution.

The binding energy is not significantly smaller in the 5^{12} cage (−6.9 kcal mol^{-1}) than in the (larger) $5^{12}6^2$ cage (−7.0 kcal mol^{-1}). The larger ΔE_{disp} in the dodecahedral cage is apparently a consequence of the smaller void: the average C–O distance is 3.88 Å in $CH_4@(H_2O)_{20}$ and 4.27 Å in $CH_4@(H_2O)_{24}$. From solid-state NMR spectra, a population ratio of 0.916 has been determined for the occupation of 5^{12} *vs.* $5^{12}6^2$ positions.[164] The relative energy difference of methane residing in one of the two cavities is therefore below 1 kcal mol^{-1}, in agreement with our data. Furthermore, the energy change of the water network after optimization of the empty cage ($-\Delta E_{def}$) is negligibly small for

Table 6 Energies (B97-D/TZV(2d,2p)) of methane hydrate models in kcal mol^{-1}

	$CH_4@(H_2O)_{20}$	$CH_4@(H_2O)_{24}$
ΔE	−6.9	−7.0
ΔE_{disp}	−10.9	−7.4
$\Delta E_{def}(H_2O)$ [a]	0.07	0.09

[a] Difference between the energies between empty water cages in the complex geometry after optimization ($E_{cpx} - E_{opt}$)

both water cages. One can conclude that the inclusion of CH_4 neither significantly distorts the cluster structure, nor influences the stability of the water cages. Similar binding energies were reported at the MP2 level (−7 kcal mol^{-1}), but only a double-zeta basis set without BSSE correction was used.[165] It is obvious from ΔE and ΔE_{disp} that DFT calculations without the dispersion correction would incorrectly predict an unbound CH_4.

4 Conclusions

The present work and the developments of dispersion corrections to Kohn–Sham density functional theory in particular adheres to the general ambition of making quantum chemical methods applicable to major parts of chemistry. We have presented several examples from organic chemistry, supramolecular chemistry, and biochemistry where non-covalent interactions are a very important issue. We have pointed out that an accurate description of the inter- and intramolecular interactions requires a balanced treatment of all basic physical processes, *i.e.*, exchange-repulsion, electrostatics, and dispersion. The inherent problem of current KS-DFT—still the most promising quantum chemical method for large systems—is the description of the dispersive (van der Waals) part, and remedies for this problem have been proposed by many groups. Our simple approach to add damped $-C_6 \cdot R^{-6}$ potentials to the KS-DFT energy captures the essential physics of the problem, as has been shown by comparisons to symmetry-adapted perturbation theory analysis. All in all, the DFT-D method with B97-D or BLYP functionals yields interaction energies that are very close to the best CCSD(T) reference data, and there is convincing evidence that this also holds for the computation of molecular geometries. One of the reasons for this success is that the complicated, non-additive and system- (and orientation-) dependent exchange-repulsion, electrostatic, and induction effects are very accurately described by the current density functionals. Dispersion forces on the other hand have a much more isotropic, system-independent (additive) character, and can thus be described with a rather simple, classical ansatz.

Currently we see little reason to make the model more complicated, *e.g.* by inclusion of higher-order terms in the dispersion correction. This would in any case require more accurate CCSD(T) reference data that would currently be too computationally demanding for most of the tested complexes. In the area of very large unsaturated systems such as carbon nanotubes or fullerene aggregates, our dispersion correction will very likely underestimate the interactions due to the large (system-dependent) polarizabilities of the fragments.

In summary, we have tried to show with some representative examples how important dispersion effects in chemistry are. These often so-called 'weak interactions' are intuitively underestimated by most chemists, but may sum up to significant relative contributions, such that their neglect can lead to qualitatively wrong conclusions. This will become more and more important as the size of the experimentally investigated molecules increases, because dispersion effects have a much longer range than the competing exchange repulsion. We hope that this work (and in particular the results of the energy decomposition analysis) will contribute to a better understanding of non-covalent interactions in large molecules. In addition, it might give experimentalists (and also theoreticians) a reality check on the importance of dispersion, and not just for density functional theory.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft in the framework of the SFB 424 ("Molekulare Orientierung als Funktionskriterium in Chemischen Systemen").

References

1 J.-M. Lehn, *Supramolecular Chemistry. Concepts and Perspectives*, VCH, Weinheim, 1995.
2 E. A. Meyer, R. K. Castellano and F. Diederich, *Angew. Chem., Int. Ed.*, 2003, **42**, 1210–1250.
3 N. Kannan and S. Vishveshwara, *Protein Eng.*, 2000, **13**, 753–761.
4 J. L. Atwood, J. E. D. Davies, D. D. MacNicol, F. Vögtle and K. S. Suslick, *Comprehensive Supramolecular Chemistry*, Elsevier, Oxford, 1996.
5 P. Hobza and J. Šponer, *Chem. Rev.*, 1999, **99**, 3247–3276.
6 K. Müller-Dethlefs and P. Hobza, *Chem. Rev.*, 2000, **100**, 143–168.
7 A. J. Stone, *The Theory of Intermolecular Forces*, Oxford University Press, Oxford, 1997.
8 C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.
9 D. Cremer, 'Møller–Plesset Perturbation Theory', in *Encyclopedia of Computational Chemistry*, vol. 3, pp. 1706–1735, ed. P. von Ragué-Schleyer, J. Wiley, New York, 1998.
10 T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, J. Wiley, New York, 2000.
11 M. O. Sinnokrot and C. D. Sherrill, *J. Phys. Chem. A*, 2006, **110**, 10656–10668.
12 S. Tsuzuki, 'Interactions with Aromatic Rings' in *Structure and Bonding*, Springer, Berlin/Heidelberg, 2005, vol. 115, pp. 149–193.
13 J. Šponer and P. Hobza, 'DNA Bases and Base Pairs: *Ab Initio* Calculations', in *Encyclopedia of Computational Chemistry*, ed. P. von Ragué Schleyer, Wiley, New York, 2001, vol. 1, pp. 777–789.
14 M. Nishio, *CrystEngComm*, 2004, **6**, 130–158.
15 W. Kutzelnigg, *Einführung in die Theoretische Chemie: Band 2, Die chemische Bindung*, Verlag Chemie, Weinheim, 1978.
16 M. Suzuki, N. Amano, J. Kakinuma and M. Tateno, *J. Mol. Biol.*, 1997, **274**, 421–435.
17 D. H. Mathews, J. Sabina, M. Zuker and D. H. Turner, *J. Mol. Biol.*, 1999, **288**, 911–940.
18 S. Bommarito, N. Peyret and J. SantaLucia, Jr, *Nucleic Acids Res.*, 2000, **28**, 1929–1934.
19 W. A. Denny and B. Baguley, *Molecular Aspects of Anticancer Drug–DNA Interactions*, ed. S. Neidle and M. Warning, Macmillan, London, 1984.
20 C. Diedrich, A. Lüchow and S. Grimme, *J. Chem. Phys.*, 2005, **123**, 184106.
21 J. G. Hill, J. A. Platts and H.-J. Werner, *Phys. Chem. Chem. Phys.*, 2006, **8**, 4072–4078.
22 Y. C. Park and J. S. Lee, *J. Phys. Chem. A*, 2006, **110**, 5091–5095.
23 R. Podeszwa, R. Bukowski and K. Szalewicz, *J. Phys. Chem. A*, 2006, **110**, 10345–10354.
24 Y. Jung and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2831–2840.
25 W. Kohn and L. Sham, *J. Phys. Rev. A*, 1965, **140**, 1133–1138.
26 W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, New York, 2001.
27 R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford, 1989.
28 S. Kristyan and P. Pulay, *Chem. Phys. Lett.*, 1994, **229**, 175–180.
29 P. Hobza, J. Sponer and T. Reschel, *J. Comput. Chem.*, 1995, **16**, 1315–1325.
30 J. M. Pérez-Jordá and A. D. Becke, *Chem. Phys. Lett.*, 1995, **233**, 134–137.
31 S. Grimme, *J. Comput. Chem.*, 2004, **25**, 1463–1473.
32 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
33 B. Zou, K. Dreger, C. Mück-Lichtenfeld, S. Grimme, H. J. Schäfer, H. Fuchs and L. Chi, *Langmuir*, 2005, **21**, 1364–1370.
34 M. Piacenza and S. Grimme, *ChemPhysChem*, 2005, **6**, 1554–1558.

35 M. Parac, M. Etinski, M. Peric and S. Grimme, *J. Chem. Theory Comput.*, 2005, **1**, 1110–1118.
36 M. Piacenza and S. Grimme, *J. Am. Chem. Soc.*, 2005, **127**, 14841–14848.
37 J. Antony and S. Grimme, *Phys. Chem. Chem. Phys.*, 2006, **8**, 5287–5293.
38 P. L. A. Popelier, 'Quantum Chemical Topology: On Bonds and Potentials', in *Structure and Bonding*, Springer: Berlin/Heidelberg, 2005, vol. 115, pp. 1–56, .
39 B. Jeziorski and K. Szalewicz, 'Intermolecular Interactions by Perturbation Theory', in *Encyclopedia of Computational Chemistry*, ed. P. von Rague-Schleyer, J. Wiley, New York, 1998, vol. 2, pp. 1376–1398.
40 C. J. Cramer, *Essentials of Compuational Chemistry*, J. Wiley, New York, 2002, p. 50.
41 H. L. Williams and C. F. Chabalowski, *J. Phys. Chem. A*, 2001, **105**, 646–659.
42 A. Heßelmann and G. Jansen, *Chem. Phys. Lett.*, 2002, **357**, 464–470.
43 A. Heßelmann and G. Jansen, *Chem. Phys. Lett.*, 2002, **362**, 319–325.
44 A. D. Mackerell, Jr, *J. Comput. Chem.*, 2004, **25**, 1584–1604.
45 F. London, *Z. Phys.*, 1930, **63**, 245–279.
46 N. A. Benedek, I. K. Snook, M. D. Towler and R. J. Needs, *J. Chem. Phys.*, 2006, **125**, 104302.
47 J. Černý and P. Hobza, *Phys. Chem. Chem. Phys.*, 2005, **7**, 1624–1626.
48 S. M. Cybulski and C. E. Seversen, *J. Chem. Phys.*, 2005, **122**, 014117.
49 T. von Mourik and R. J. Gdanitz, *J. Chem. Phys.*, 2002, **116**, 9620–9623.
50 M. J. Allen and D. J. Tozer, *J. Chem. Phys.*, 2002, **117**, 11113–11120.
51 J. M. Pérez-Jordá, E. San-Fabián and A. J. Pérez-Jiménez, *J. Chem. Phys.*, 1999, **110**, 1916–1920.
52 P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864–871.
53 M. P. Waller, A. Robertazzi, J. A. Platts, D. E. Hibbs and P. A. Williams, *J. Comput. Chem.*, 2006, **27**, 491–504.
54 X. Xu and W. A. Goddard, III, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 2673–2677.
55 T. R. Walsh, *Phys. Chem. Chem. Phys.*, 2005, **7**, 443–451.
56 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Theory Comput.*, 2006, **2**, 364–382.
57 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 6908–6918.
58 I. V. Schweigert, V. F. Lotrich and R.J. Bartlett, *J. Chem. Phys.*, 2006, **125**, 104108.
59 O. Gritsenko and E. J. Baerends, *J. Chem. Phys.*, 2006, **124**, 054115.
60 S. Grimme, *J. Chem. Phys.*, 2006, **124**, 034108.
61 A. Puzder, M. Dion and D. C. Langreth, *J. Chem. Phys.*, 2006, **124**, 164105.
62 T. Thonhauser, A. Puzder and D. C. Langreth, *J. Chem. Phys.*, 2006, **124**, 164106.
63 I. C. Gerber and J. G. Ángyán, *Chem. Phys. Lett.*, 2005, **416**, 370–375.
64 J. G. Ángyán, I. C. Gerber, A. Savin and J. Toulouse, *Phys. Rev. A*, 2005, **72**, 012510.
65 M. A. Basanta, Y. J. Dappe, J. Ortega and F. Flores, *Europhys. Lett.*, 2005, **70**, 355–361.
66 F. Furche and T. Van Voorhis, *J. Chem. Phys.*, 2005, **122**, 164106.
67 T. Sato, T. Tsuneda and K. Hirao, *J. Chem. Phys.*, 2005, **123**, 104307.
68 E. Hult, H. Rydberg, B. I. Lundqvist and D. C. Langreth, *Phys. Rev. B: Condens. Matter*, 1999, **59**, 4708–4713.
69 W. Kohn, Y. Meir and D. E. Makarov, *Phys. Rev. Lett.*, 1998, **80**, 4153–4156.
70 Y. Andersson, D. C. Langreth and B. I. Lundqvist, *Phys. Rev. Lett.*, 1996, **76**, 102–105.
71 E. Hult, Y. Andersson, B. I. Lundqvist and D. C. Langreth, *Phys. Rev. Lett.*, 1996, **77**, 2029–2032.
72 B. I. Lundqvist, Y. Andersson, H. Shao, S. Chan and D. C. Langreth, *Int. J. Quantum Chem.*, 1995, **56**, 247–255.
73 M. Elstner, P. Hobza, T. Frauenheim, S. Suhai and E. Kaxiras, *J. Chem. Phys.*, 2001, **114**, 5149–5155.
74 X. Wu, M. C. Vargas, S. Nayak, V. Lotrich and G. Scoles, *J. Chem. Phys.*, 2001, **115**, 8748–8757.
75 Q. Wu and W. Yang, *J. Chem. Phys.*, 2002, **116**, 515–524.
76 U. Zimmerli, M. Parrinello and P. Koumoutsakos, *J. Chem. Phys.*, 2004, **120**, 2693–2699.
77 A. D. Becke and E. R. Johnson, *J. Chem. Phys.*, 2005, **122**, 154104.
78 A. D. Becke and E. R. Johnson, *J. Chem. Phys.*, 2005, **123**, 154101.
79 A. D. Becke and E. R. Johnson, *J. Chem. Phys.*, 2006, **124**, 014104.
80 S. Grimme, *Angew. Chem.*, 2006, **118**, 4571–4575.
81 E. R. Johnson and A. D. Becke, *J. Chem. Phys.*, 2005, **123**, 024101.
82 E. R. Johnson and A. D. Becke, *J. Chem. Phys.*, 2006, **124**, 174104.
83 J. Hepburn, G. Scoles and R. Penco, *Chem. Phys. Lett.*, 1975, **36**, 451–456.
84 R. Ahlrichs, R. Penco and G. Scoles, *Chem. Phys.*, 1977, **19**, 119–130.
85 K. T. Tang and J. P. Toennies, *J. Chem. Phys.*, 1984, **80**, 3726–3741.
86 C. Douketis, G. Scoles, S. Marchetti, M. Zen and A.J. Thakkar, *J. Chem. Phys.*, 1982, **76**, 3057–3063.
87 I. Hyla-Kryspin, G. Haufe and S. Grimme, *Chem.–Eur. J.*, 2004, **10**, 3411–3422.
88 X. Xu and W. A. Goddard, III, *J. Phys. Chem. A*, 2004, **108**, 2305–2313.
89 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2006, **110**, 5121–5129.
90 Y. Zhao and D. G. Truhlar, *Phys. Chem. Chem. Phys.*, 2005, **7**, 2701–2705.
91 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 4209–4212.
92 V. N. Staroverov, G. E. Scuseria, J. Tao and J. P. Perdew, *J. Chem. Phys.*, 2003, **119**, 12129–12137.
93 J. Tao and J. P. Perdew, *J. Chem. Phys.*, 2005, **122**, 114102.
94 A. Ruzsinszky, J. P. Perdew and G. I. Csonka, *J. Phys. Chem. A*, 2005, **109**, 11015–11021.
95 C. Adamo and V. Barone, *J. Chem. Phys.*, 1998, **108**, 664–675.
96 D. C. Langreth, M. Dion, H. Rydberg, E. Schröder, P. Hyldgaard and B. I. Lundqvist, *Int. J. Quantum Chem.*, 2005, **101**, 599–610.
97 M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist, *Phys. Rev. Lett.*, 2004, **92**, 246401.
98 S. D. Chakarova and E. Schröder, *J. Chem. Phys.*, 2005, **122**, 054102.
99 J. Kleis and E. Schröder, *J. Chem. Phys.*, 2005, **122**, 164902.
100 E. Goll, H.-J. Werner and H. Stoll, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3917–3923.
101 S. M. Cybulski, T. M. Bledson and R. R. Toczylowski, *J. Chem. Phys.*, 2002, **116**, 11039–11040.
102 O. A. von Lilienfeld, I. Tavernelli, U. Röthlisberger and D. Sebastiani, *Phys. Rev. B: Condens. Matter*, 2005, **71**, 195119.
103 O. A. von Lilienfeld, I. Tavernelli, U. Röthlisberger and D. Sebastiani, *J. Chem. Phys.*, 2005, **122**, 014113.
104 O. A. von Lilienfeld, I. Tavernelli, U. Röthlisberger and D. Sebastiani, *Phys. Rev. Lett.*, 2004, **93**, 153004.
105 O. A. von Lilienfeld and D. Andrienko, *J. Chem. Phys.*, 2006, **124**, 054307.
106 G. Cinacchi, *J. Chem. Phys.*, 2006, **125**, 057101.
107 T. Tkatchenko and O. A. von Lilienfeld, *Phys. Rev. B: Condens. Matter*, 2006, **73**, 153406.
108 C. Tuma and J. Sauer, *Chem. Phys. Lett.*, 2004, **387**, 388–394.
109 C. Tuma and J. Sauer, *Angew. Chem., Int. Ed.*, 2005, **44**, 4769–4771.
110 C. Tuma and J. Sauer, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3955–3965.
111 S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma and M. J. Frisch, *J. Mol. Struct. (THEOCHEM)*, 1999, **461–462**, 1–21.
112 E. Clementi and G. Corongiu, *J. Phys. Chem. A*, 2001, **105**, 10379–10383.
113 C. Gonzalez and E. C. Lim, *J. Phys. Chem. A*, 2003, **107**, 10105–10110.
114 T. H. Dunning, Jr, *J. Chem. Phys.*, 1989, **90**, 1007–1023.
115 H. B. Jansen and P. Ros, *Chem. Phys. Lett.*, 1969, **3**, 140–143.
116 S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–566.
117 A. Schäfer, C. Huber and R. Ahlrichs, *J. Chem. Phys.*, 1994, **100**, 5829–5835.
118 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
119 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter*, 1988, **37**, 785–789.
120 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
121 A. D. Becke, *J. Chem. Phys.*, 1997, **107**, 8554–8560.
122 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
123 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
124 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
125 R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165–169.
126 *TURBOMOLE, version 5.7*, R. Ahlrichs, M. Bär, H.-P. Baron, R. Bauernschmitt, S. Böcker, N. Crawford, P. Deglmann, M. Ehrig, K. Eichkorn, S. Elliott *et al.*, Universität Karlsruhe, 2003, http://www.turbomole.com.
127 K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chem. Phys. Lett.*, 1995, **240**, 283–289.

128 F. Weigend and M. Häser, *Theor. Chem. Acc.*, 1997, **97**, 331–340.
129 F. Weigend, A. Köhn and C. Hättig, *J. Chem. Phys.*, 2002, **116**, 3175–3183.
130 K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theor. Chem. Acc.*, 1997, **97**, 119–124.
131 The basis sets are available from the TURBOMOLE homepage (http://www.turbomole.com) *via* the FTP Server button (in the subdirectories basen, jbasen, and cbasen).
132 H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. Deegan, O. A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni and T. Thorsteinsson, *MOLPRO, version 2006.1:, A package of* ab initio *programs*, http://www.molpro.net.
133 A. Heßelmann, G. Jansen and M. Schütz, *J. Chem. Phys.*, 2005, **122**, 014103.
134 K. Morokuma, *J. Chem. Phys.*, 1971, **55**, 1236–1244.
135 K. Kitaura and K. Morokuma, *Int. J. Quantum Chem.*, 1976, **10**, 325–340.
136 F. M. Bickelhaupt and E. J. Baerends, 'Kohn–Sham Density Functional Theory: Predicting and Understanding Chemistry' in *Reviews in Computational Chemistry*, ed. K. B. Lipkowitz and D. B. Boyd, Wiley-VCH, New York, 2000, vol. 15, pp. 1–86.
137 M. Swart, C. Fonseca Guerra and F. M. Bickelhaupt, *J. Am. Chem. Soc.*, 2004, **126**, 16718–16719.
138 M. O. Sinnokrot and C. D. Sherrill, *J. Am. Chem. Soc.*, 2003, **126**, 7690–7697.
139 P. Jurecka, J. Sponer, J. Cerny and P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985–1993.
140 P. Mignon, S. Loverix, F. De Proft and P. Geerlings, *J. Phys. Chem. A*, 2004, **108**, 6038–6044.
141 P. Mignon, S. Loverix, J. Steyaert and P. Geerlings, *Nucleic Acids Res.*, 2005, **33**, 1779–1789.
142 C. A. Hunter and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1990, **112**, 5525–5534.
143 S. Grimme, *Chem.–Eur. J.*, 2004, **10**, 3423–3429.
144 S. Tsuzuki, K. Honda, T. Uchimaru and M. Mikami, *J. Chem. Phys.*, 2006, **124**, 114304.
145 B. W. Gung, X. Xue and H. J. Reich, *J. Org. Chem.*, 2005, **70**, 3641–3644.
146 A. Klamt and G. Schürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799–805.
147 Q. Wu and W. Yang, *J. Chem. Phys.*, 2002, **116**, 515–524.
148 J. N. S. Evans, G. Burton, P. E. Mackenzie and N. E. Scott, *Biochemistry*, 1986, **25**, 905–912.
149 N. L. Allinger, Y. H. Yuh and J.-H. Lii, *J. Am. Chem. Soc.*, 1989, **111**, 8551–8566.
150 J. Fritzsche, *J. Prakt. Chem. (Leipzig)*, 1867, **101**, 333–343.
151 S. Grimme, C. Diedrich and M. Korth, *Angew. Chem.*, 2006, **118**, 641–645.
152 S. Grimme, *J. Chem. Phys.*, 2003, **118**, 9095–9102.
153 S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami and K. Tanabe, *J. Am. Chem. Soc.*, 2002, **124**, 104–112.
154 M. O. Sinnokrot, E. F. Valeev and C. D. Sherrill, *J. Am. Chem. Soc.*, 2002, **124**, 10887–10893.
155 S. Tsuzuki, T. Uchimaru and K. Tanabe, *J. Phys. Chem. A*, 1997, **102**, 740–743.
156 M. A. Petrukhina, Y. Sevryugina, A. Y. Rogachev, E. A. Jackson and L. T. Scott, *Organometallics*, 2006, **25**, 5492–5495.
157 S. Yao, U. Beginn, T. Gress, M. Lysetska and F. Würthner, *J. Am. Chem. Soc.*, 2004, **126**, 8336–8348.
158 M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902–3909.
159 M. Rapacioli, F. Calvo, F. Spiegelman, C. Joblin and D. J. Wales, *J. Phys. Chem. A*, 2005, **109**, 2487–2497.
160 U. Rösch, S. Yao, R. Wortmann and F. Würthner, *Angew. Chem., Int. Ed.*, 2006, **45**, 7026–7030.
161 G. Vidali, G. Ihm, H. Kim and M. W. Cole, *Surf. Sci. Rep.*, 1991, **12**, 133–181.
162 S. E. Weber, S. Talapatra, C. Journet, A. Zambano and A. D. Migone, *Phys. Rev. B: Condens. Matter*, 2000, **61**, 13150–13154.
163 W. F. Claussen, *J. Chem. Phys.*, 1951, **19**, 1425–1426.
164 J. Ripmeester and C. I. Ratcliffe, *J. Phys. Chem.*, 1988, **92**, 337–339.
165 A. Khan, *J. Chem. Phys.*, 1999, **110**, 11884–11889.

