

Effect of the Damping Function in Dispersion Corrected Density Functional Theory

STEFAN GRIMME,¹ STEPHAN EHRLICH,¹ LARS GOERIGK^{1,2}

¹Theoretische Organische Chemie, Organisch-Chemisches Institut der Universität Münster, Corrensstraße 40

²NRW Graduate School of Chemistry, Wilhelm-Klemm-Straße 10, D-48149 Münster, Germany

Received 4 November 2010; Revised 22 December 2010; Accepted 28 December 2010

DOI 10.1002/jcc.21759

Published online 1 March 2011 in Wiley Online Library (wileyonlinelibrary.com).

Abstract: It is shown by an extensive benchmark on molecular energy data that the mathematical form of the damping function in DFT-D methods has only a minor impact on the quality of the results. For 12 different functionals, a standard “zero-damping” formula and rational damping to finite values for small interatomic distances according to Becke and Johnson (BJ-damping) has been tested. The same (DFT-D3) scheme for the computation of the dispersion coefficients is used. The BJ-damping requires one fit parameter more for each functional (three instead of two) but has the advantage of avoiding repulsive interatomic forces at shorter distances. With BJ-damping better results for nonbonded distances and more clear effects of intramolecular dispersion in four representative molecular structures are found. For the noncovalently-bonded structures in the S22 set, both schemes lead to very similar intermolecular distances. For noncovalent interaction energies BJ-damping performs slightly better but both variants can be recommended in general. The exception to this is Hartree-Fock that can be recommended only in the BJ-variant and which is then close to the accuracy of corrected GGAs for non-covalent interactions. According to the thermodynamic benchmarks BJ-damping is more accurate especially for medium-range electron correlation problems and only small and practically insignificant double-counting effects are observed. It seems to provide a physically correct short-range behavior of correlation/dispersion even with unmodified standard functionals. In any case, the differences between the two methods are much smaller than the overall dispersion effect and often also smaller than the influence of the underlying density functional.

© 2011 Wiley Periodicals, Inc. J Comput Chem 32: 1456–1465, 2011

Key words: dispersion energy; density functional theory; noncovalent interactions; van der Waals complexes

Introduction

The development of approximate density functional theory (DFT) approaches that accurately model the physically and chemically very important London dispersion interactions^{1,2} is a very active field of research (for a recent review see ref. 3). Inclusion of these interactions in theoretical simulations of weakly bonded (van der Waals, vdW) systems⁴ is indispensable to reach so-called chemical accuracy, but also influences the accuracy of theoretical thermodynamics.^{5–7} Spectacular improvements by using dispersion corrections in DFT are also found in solid state applications (see e.g. ref. 8 for various examples).

One of the most successful and widely used methods in the field is DFT-D in which a damped, atom-pair wise potential is added to a standard Kohn-Sham DFT result. The idea (a kind of QM/MM hybrid scheme) goes back to the 1970s in the context of Hartree-Fock theory^{9,10} or special local density approximations.¹¹ The method has been forgotten for almost 30 years and was rediscovered about 10 years ago^{12–14} as the DFT problems for dispersion dominated interactions became more evident.

Meanwhile, many modifications of the DFT-D approach have been published. All are based on an atom pair-wise additive treatment of the dispersion energy (for extensions to include three-body non-additive dispersion effects see refs. 15, 16). The general form for the dispersion energy as discussed here is

$$E_{\text{disp}}^{\text{DFT-D}} = -\frac{1}{2} \sum_{A \neq B} \sum_{n=6,8,10,\dots} s_n \frac{C_n^{\text{AB}}}{R_{\text{AB}}^n} f_{\text{damp}}(R_{\text{AB}}). \quad (1)$$

Here, the sum is over all atom pairs in the system, C_n^{AB} denotes the averaged (isotropic) n th-order dispersion coefficient (orders $n = 6, 8, 10, \dots$) for atom pair AB, and R_{AB} is their internuclear distance. Global (functional dependent) scaling factors s_n can be used to adjust

Correspondence to: S. Grimme; e-mail: grimmes@uni-muenster.de

Contract/grant sponsor: The Deutsche Forschungsgemeinschaft (SFB 858)

Contract/grant sponsor: Fonds der chemischen Industrie

the correction to the repulsive behavior of the chosen exchange-correlation density functional (DF).¹⁷ One of the key ingredients in all DFT-D methods is the damping function f_{damp} . It determines the short-range behavior of the dispersion correction and is needed to avoid near-singularities for small R_{AB} and double-counting effects of electron correlation at intermediate distances. For a discussion of general damping functions see ref. 18. Typical expressions are¹⁹

$$f_{\text{damp}}(R_{\text{AB}}) = \frac{1}{1 + 6(R_{\text{AB}}/(s_{r,n}R_0^{\text{AB}}))^{-\gamma}}, \quad (2)$$

or^{14,20}

$$f_{\text{damp}}(R_{\text{AB}}) = \frac{1}{1 + e^{-\gamma(R_{\text{AB}}/s_{r,n}R_0^{\text{AB}} - 1)}}, \quad (3)$$

where R_0^{AB} is a cut-off radius for atom pair AB, $s_{r,n}$ is a functional dependent (global) radii scaling factor (as first introduced in ref. 21) and γ is a parameter that determines the steepness of the function for small R_{AB} . For the R_0^{AB} values, often (averaged) empirical atomic vdW-radii are used (for an *ab initio* approach to determine pair-specific values see ref. 15). Another (more complicated) damping function that is used in some DFT-D methods is that of Tang and Toennies.²²

It is generally assumed that the damping function represents an important (and empirical) element in the entire approach and that the specific choice of its form (including the parameters) strongly influences the results. The damping function was criticized²³ as one of the weak points of DFT-D. On the contrary, we have observed in the last 7 years of experience with our DFT-D schemes^{15,17,20} that the damping function has only minor impact on the results if properly fitted, although this has never been quantified. This issue is not only of technical nature but related to the fundamental question how the dispersion energy (a form of electron correlation effect) behaves asymptotically in the $R_{\text{AB}} \rightarrow 0$ (united atom) limit. While the above given formulas (and thus E_{disp}) approach zero for $R_{\text{AB}} \rightarrow 0$, Becke and Johnson (BJ)^{24–26} propose rational damping in the following form ($n \geq 6$)

$$E_{\text{disp}} = -\frac{1}{2} \sum_{A \neq B} \frac{C_n^{\text{AB}}}{R_{\text{AB}}^n + \text{const.}}. \quad (4)$$

This so-called BJ-damping leads to a constant contribution of E_{disp} to the total correlation energy from each spatially close pair of atoms (i.e., directly bonded). According to a convergent multipole-expansion of the dispersion energy as introduced by Koide,²⁷ the dispersion energy should approach a constant (finite) value for $R_{\text{AB}} \rightarrow 0$ (which e.g., for $2\text{H} \rightarrow \text{He}$ amounts to about -8 mE_h for the dipole-dipole term²⁷). Thus, although BJ-damping seems theoretically to be more justified than zero-damping, it changes the thermochemical description of the underlying DF for “normal” bonding situations. As a result, adjustment of standard correlation functionals might be required. The difference between the two “damping-philosophies” is illustrated schematically in Figure 1. Note that a disadvantage of the zero-damping approach is that at small and

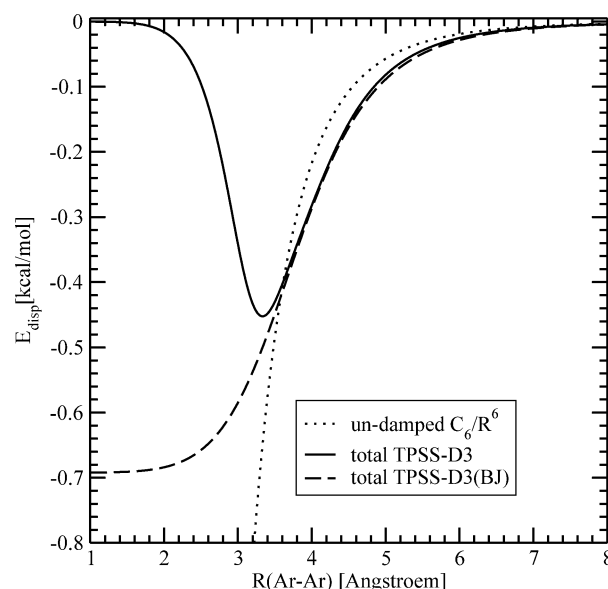


Figure 1. Dispersion correction for two argon atoms (including sixth- and eighth-order terms, dispersion coefficients from ref. 15) with the zero- and BJ-damping methods in comparison with the un-damped $-C_6 \cdot R^{-6}$ term.

medium distances (“left” of the minimum in Fig. 1) the atoms experience a repulsive force which may lead (counter-intuitively) to longer interatomic distances with dispersion correction than without.

The major aim of this work is to investigate thoroughly the effect of the damping function on the results for extensive thermochemical benchmarks using standard functionals. In particular, we want to answer the question if the form of the damping really has a significant impact. Note, that zero- and BJ-damping differ as much as possible so that our conclusions are readily generalizable. Second, our investigation might give a hint if the physically more sound BJ-damping improves the description of correlation effects in “normal” chemical reactions and if adjustments of correlation functionals are necessary. To the best of our knowledge, this aspect of BJ-damping has never been thoroughly studied for standard DFs (for applications of the BJ-damping scheme with non-standard DFs see e.g. refs. 28, 29). We also here for the first time test the relatively new DFT-D3 model (in both variants) for the computation of intermolecular complex geometries (S22 set³⁰) and nonbonded distances in molecules. In particular, we shine some light on the importance of intramolecular dispersion effects for structures with unusually close-lying atoms (below vdW distance). As examples, C_2Br_6 , S_8^{2+} , (N,N-dimethylaminoxy)trifluorosilane (DTFS), and [2.2]paracyclophane (22PC) are investigated (for molecular structures see Fig. 2).

Methods

Currently, the most widely used DFT-D method is our version from 2006²⁰ (now termed DFT-D2), which represents an update of DFT-D1 from 2004.¹⁷ The method has recently been refined regarding

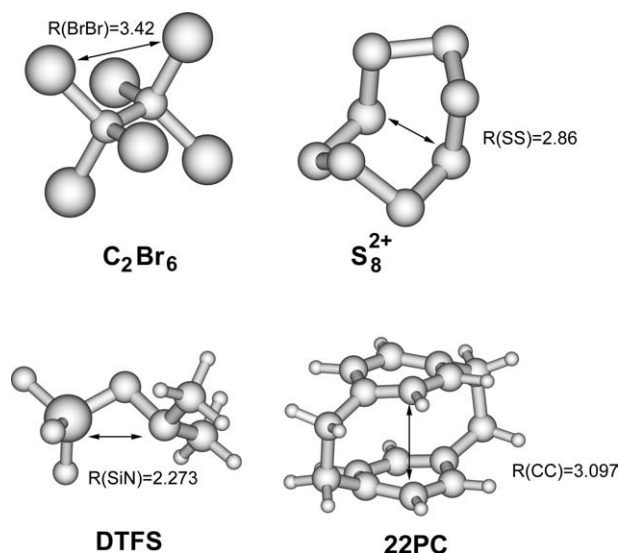


Figure 2. Investigated molecular structures with short nonbonded interactions. The distances considered (experimental values in Å) are indicated by an arrow.

higher accuracy, broader range of applicability, and less empiricism (called DFT-D3¹⁵). The main new ingredients are atom-pair wise specific dispersion coefficients and a new set of cut-off radii, both computed from first principles. The coefficients for eight-order dispersion terms are computed using established recursion relations from the C_6 coefficients. System (geometry) dependent information is used for the first time in a non-electronic DFT-D type approach by employing the new concept of fractional coordination numbers (for another geometry-dependent DFT-D method that, however, requires electronic information see ref. 31). This allows one to distinguish in a differentiable way the different hybridization states of atoms in molecules. The method has been introduced using eq. (2) for damping with $n = 6, 8$. It only requires adjustment of two global parameters for each density functional, is asymptotically exact ($s_6 = 1$) for a gas of weakly interacting neutral atoms and easily allows the computation of atomic forces. Accurate dispersion coefficients and cut-off radii are available for all elements up to $Z = 94$. This revised DFT-D3 method can be used analogously with BJ-damping in which the dispersion energy is given by

$$E_{\text{disp}}^{\text{D3(BJ)}} = -\frac{1}{2} \sum_{A \neq B} s_6 \frac{C_6^{\text{AB}}}{R_{\text{AB}}^6 + [f(R_{\text{AB}}^0)]^6} + s_8 \frac{C_8^{\text{AB}}}{R_{\text{AB}}^8 + [f(R_{\text{AB}}^0)]^8} \quad (5)$$

with

$$f(R_{\text{AB}}^0) = a_1 R_{\text{AB}}^0 + a_2 \quad (6)$$

Here, the sum runs over all atom pair in the system and a_1 and a_2 are free fit parameters introduced by BJ. We investigated two choices for R_{AB}^0 . The first uses the same values as in DFT-D3 while the second employs the following relation²⁶

$$R_{\text{AB}}^0 = \sqrt{\frac{C_8^{\text{AB}}}{C_6^{\text{AB}}}}. \quad (7)$$

After some testing, it turns out that the second variant [eq. (7)] performs slightly better in the BJ-damping approach and it will therefore exclusively be discussed here (for zero-damping the original R_{AB}^0 values are taken). Because in DFT-D3 the higher-ranked C_n values are computed recursively from C_6 values,¹⁵ eq. (7) reduces here to a geometric mean of (tabulated) atomic $\frac{\langle r^4 \rangle}{\langle r^2 \rangle}$ expectation values. Also as in DFT-D3, s_6 is set to unity for GGA and hybrid functionals and s_8 is used to adapt the correction to the repulsive character of the short/medium-range behavior of the exchange correlation functional. Note that if γ is fixed in eq. (2) as usual, the number of free fit parameters for each DF is three in DFT-D3(BJ) (s_8 , a_1 , and a_2) but only two in DFT-D3 ($s_{r,6}$ and s_8). In comparison, s_8 and $s_{r,6}/a_1$ have the same physical meaning in DFT-D3 and DFT-D3(BJ), whereas a_2 is found to be necessary for “fine-tuning” of the BJ-damping. Equation (5) and the corresponding derivatives have been implemented in the freely available `df-t-d3` code.³²

In the following, we will use the term “DFT-D3” (or with appended “-D3” to a functional name) to denote the original version with zero-damping and dub the variant employing BJ-damping as “DFT-D3(BJ)” [(or “-D3(BJ)”).

Technical Details of the Calculations

The DFT calculations have been performed with the TURBOMOLE 5.9³³ programs. If not mentioned otherwise the results refer to calculations using the very large def2-QZVP Gaussian AO basis^{34,35} and are based for open-shell systems on unrestricted (symmetry broken) KS-wavefunctions. The def2-QZVP sets are, for example, for the first-row atoms of [7s4p] quality and are augmented with polarization functions typically taken from Dunning quadruple-zeta sets leading finally to [7s4p3d2f1g] sets. They provide results quite close to the KS-limit for many properties and are also essentially free of basis set superposition error (BSSE). The exponents of the outmost primitive Gaussian functions in these sets can be considered as being semi-diffuse so that def2-QZVP can be used for large systems without risking linear dependencies and concomitantly yield acceptable polarizabilities with most GGA functionals. Counterpoise-corrections are not applied. For the heavier elements ($Z > 36$) small core effective pseudo-potentials of the Stuttgart/Cologne groups³⁶ were used. Large numerical quadrature grids of size $m4$ or $m5$ (in TURBOMOLE notation) were generally employed. To speed up the calculations, the RI-approximation for the SCF-part (RI-J or RI-JK^{37,38}) and the perturbation part of double-hybrid functionals³⁹ was used. Matching auxiliary basis functions^{40–42} were taken from the TURBOMOLE basis set library.

The three empirical parameters in the BJ-damping for each DF have been determined analogously as for DFT-D3 by a least-squares fit to (relative) reference energies in non-covalent interaction energy benchmark sets (S22, S22+, PCONE, SCONE, ACONE, CCONE, ADIM6, RG6; for details see ref. 15). Note that we use the recently revised reference energies in S22.⁴³ A total of 130 reference data values were included in the fit.

Table 1. Mean Absolute Deviation (MAD, in kcal/mol) Averaged Over 12 Density Functionals for Noncovalent Interaction Energy Benchmarks. For Comparison MP2 Values (Also Using the def2-QZVP AO Basis) are Given.

Subset	DFT-D3	DFT-D3(BJ)	MP2
All	0.37	0.36	0.41
S22	0.40	0.36	1.02
S22+	0.28	0.25	0.42
ADIM6	0.50	0.23	0.04
RG6	0.06	0.06	0.11
ACONF	0.11	0.10	0.10
SCONF	0.47	0.47	0.12
CCONF	0.49	0.61	0.24
PCONF	0.76	1.00	0.73

Results and Discussion

Benchmarks for Noncovalent Interactions

The most important question of this work is about the performance of the two different damping schemes for typical noncovalent interactions, which is the main purpose of DFT-D3. Our standard benchmark, on which DFT-D3 has originally been tested¹⁵ and that is used to determine the functional-dependent parameters, is employed first. It consists of noncovalently bound complexes (S22, S22+, ADIM6, RG6) and four intramolecular (conformational energies) subsets (PCONF, SCONF, ACONF, CCONF). According to previous experience,¹⁵ it gives a quite representative picture about the performance of various methods. A cross-validation study is also performed using the noncovalent subsets WATER27, IDISP and HEAVY28 from the GMTKN30 database⁴⁴ (see next section).

The mean absolute deviations (MADs) for the entire set (130 energies, labeled “all”) and for the subsets are given in Table 1. These data have been obtained by averaging over 12 typical DFs (BLYP,^{45,46} BP86,^{45,47,48} PBE,⁴⁹ revPBE,⁵⁰ B97-D,²⁰ rPW86PBE,^{49,51} TPSS,⁵² B3LYP^{53,54} PBE0,⁵⁵ TPSS0,⁵⁶ PW6B95,⁵⁷ B2PLYP⁵⁸). Results for individual DFs are given in Figure 3 and the optimized DFT-D3(BJ) parameters are given in Table 2. Parameters have also been determined for Hartree-Fock (HF) but these results are not included in averaged MAD values and graphics and are discussed separately. Optimized DFT-D3 parameters for rPW86PBE ($s_{r,6} = 1.224$, $s_8 = 0.901$) and rPW86PBE-D3 results are reported in this work for the first time.

Before discussing the various dispersion-corrected DFT results, it should be noted that the performance of plain DFT for noncovalent interactions is disastrous. The MADs are typically 5–10 times larger than for both versions of DFT-D3 and many systems are unbound or conformer orderings are qualitatively wrong (see also Section “Thermodynamic Benchmarks”). Thus, the discussion of the differences between DFT-D3 and DFT-D3(BJ) happens on a rather high level of accuracy (that is typically higher than that of the simplest correlated wave-function based method which is MP2, see Table 1). From inspection of the averaged results in Table 1, it is clear that both versions perform very similar and overall the MAD for DFT-D3(BJ) is slightly lower compared to DFT-D3. (0.36 vs. 0.37 kcal/mol). DFT-D3(BJ) is significantly better for S22 and the

alkane dimers, whereas the opposite holds for cysteine and peptide conformers (CCONF and PCONF).

The Figure 3 illustrates this for all sets and the S22 and PCONF subsets but now specifically for all considered functionals. No clear trend regarding the type or composition of the exchange-correlation functional in favor of one or the other damping scheme is seen. For all 130 entries only for B97-D and revPBE is DFT-D3(BJ) significantly worse than DFT-D3. The opposite holds for rPW86PBE, BLYP, and B3LYP. For the S22 subset DFT-D3(BJ) is with the

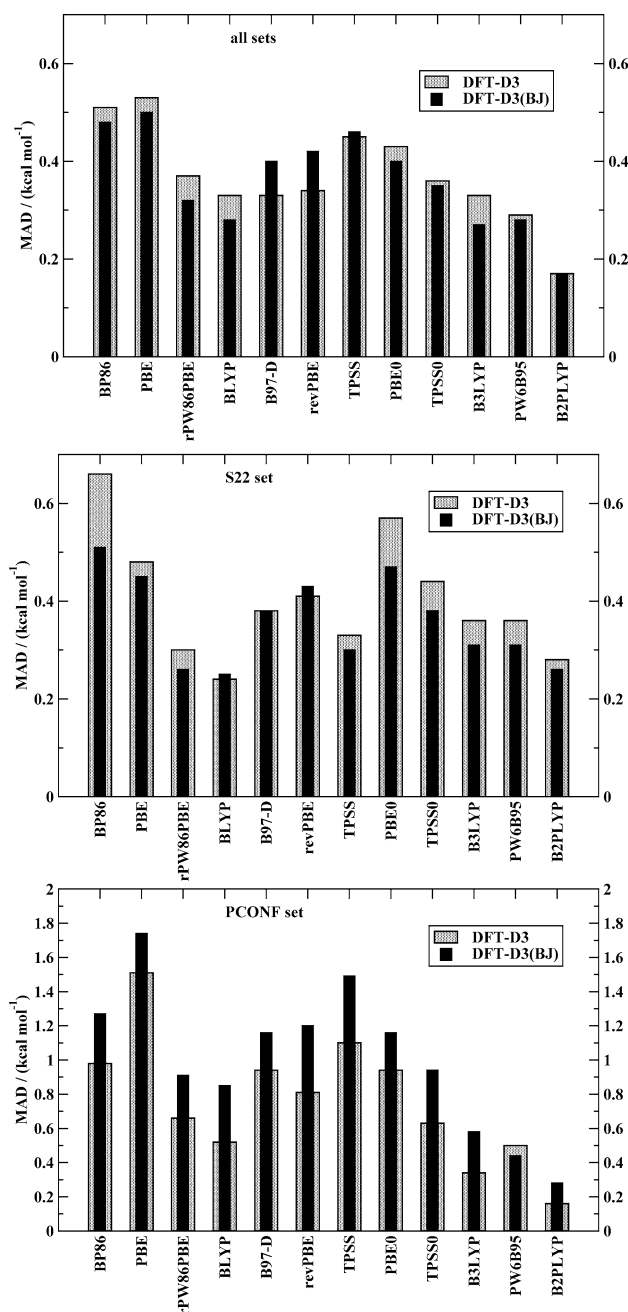
**Figure 3.** Mean absolute deviation (MAD) of 12 density functionals. Top: all 130 entries; middle: S22 subset; bottom: PCONF subset.

Table 2. Optimized DFT-D3(BJ) Parameters (in atomic units) for 12 Density Functionals and Hartree-Fock.

Functional	a_1	s_8	a_2
BP86	0.3946	3.2822	4.8516
PBE	0.4289	0.7875	4.4407
rPW86PBE	0.4613	1.3845	4.5062
BLYP	0.4298	2.6996	4.2359
B97-D	0.5545	2.2609	3.2297
revPBE	0.5238	2.3550	3.5016
TPSS	0.4535	1.9435	4.4752
PBE0	0.4145	1.2177	4.8593
TPSS0	0.3768	1.2576	4.5865
B3LYP	0.3981	1.9889	4.4211
PW6B95	0.2076	0.7257	6.3750
B2PLYP ^a	0.3451	1.0860	4.7735
HF ^b	0.3385	0.9171	2.8830

^aGlobal scaling factor of $s_6 = 0.5$ as in DFT-D3¹⁵ is used due to nonlocal correlation part. For a revised s_6 value and a general theoretical procedure for its determination for double-hybrid functionals see ref. 44.

^bHartree-Fock.

exception of revPBE and BLYP always better than DFT-D3. This is reversed for the peptide conformers in PCONF for which DFT-D3 is clearly better. At present we have no clear understanding of the “outlying” behavior of the PCONF subset but note that this benchmark is one of the most challenging in our noncovalent interaction database and very well described only by highly non-local DFs (e.g., B2PLYP). In summary, one can conclude that both versions on average yield almost the same high accuracy for typical noncovalent interactions and that the observed differences seem to reflect basic flaws of the underlying functionals. It can be expected that the performance of the two versions further diminishes when the correlation functional is specifically adjusted to the presence of the correction (which, however, seems not to be necessary).

The optimized parameters are given in Table 2. As for DFT-D3, the different “repulsiveness” of the DF is reflected mainly in the parameter s_8 which is, for example, larger for all functionals based on B88 exchange. The values for a_1 and a_2 are mostly of about the same magnitude as those optimized by Kannemann and Becke.^{29,59} DFs, which inherently include a larger fraction of “medium-range correlation” typically have smaller(larger) a_1 (a_2) parameters, which makes the correction more long-ranged and less atom-pair dependent. Note that a_1 can also become negative meaning that for “larger” atom pairs a lower $R \rightarrow 0$ asymptotic dispersion energy is obtained. Such behavior seems rather reasonable theoretically because the correlation energy linearly depends on the number of electrons. However, this is not found for the “normal” DFs considered because they intrinsically describe this region by density-dependent terms and avoid double-counting effects by $a_1 > 0$. Negative values for a_2 as reported recently⁵⁹ have been tested as well but in general do not lead to consistently accurate corrections. In passing, we note that parametrization of DFT-D3(BJ) is a bit more complicated and that for a few functionals more than one minimum on the error hypersurface with respect to the three parameters have been observed. In this context it is strongly recommended to use extended test sets for the determination and evaluation of dispersion corrections. Just

taking, for example, the S22 data set alone is clearly insufficient for this purpose.

We also tested for Hartree-Fock zero- and BJ-damping. The optimized parameters are given in Table 2. The corresponding DFT-D3 parameters $s_{r,6}$ and s_8 are 1.158 and 1.746, respectively. For the complete fit set with 130 entries, HF-D3 yields a relatively large MAD of 1.04 kcal/mol, which is clearly worse than for the worst DF (which is PBE with MAD = 0.50 kcal/mol). However, the BJ scheme very strongly improves for HF and in fact we see here the by far largest effect of the choice of the damping in our study. The MAD of HF-D3(BJ) drops by about 40% to a value of only 0.65 kcal/mol, which is in the same range as for the worst GGAs like BP86 or PBE. This indicates that completely uncorrelated (more repulsive) methods benefit most from the better description of medium-range correlation in the BJ-scheme. This is in line with the good results for rPW86PBE-D3(BJ) and BLYP-D3(BJ). Although the remaining errors in HF-D3(BJ) are not small (e.g., the MAD for the S22 set is only 0.87 kcal/mol), they are rather uniformly distributed and no particular outliers are found. In fact the method performs as expected and is not as good for, for example, hydrogen-bonded systems in which the short-ranged parts of the electrostatic and induction energies (that are much better described with DFs) are important. We suggest HF-D3(BJ) as a quantum chemical method for non-covalent interactions when self-interaction error related problems are important and other alternatives are computationally too demanding.

Thermodynamic Benchmarks

One of the further questions of this work is if the parameters a_1 , a_2 , and s_8 , as determined from a fit to noncovalent interactions, can also be used for the calculation of “normal” thermodynamic properties of molecules. For DFT-D3, this has been shown already in the original work¹⁵ and recently for alkane isodesmic reactions.^{60,61} A very convenient and un-biased way to investigate this is to use our recently proposed GMTKN30 database,⁴⁴ which is an extension of GMTKN24.⁶²

This benchmark set covers 30 subsets related to general main group thermochemistry, kinetics and non-covalent interactions. In total, it comprises 1218 single point calculations and 841 data points (relative energies) and, thus, it turned out to be ideal for the evaluation and development of DFT methods. The subsets of GMTKN30 can be divided into three major sections. These are basic properties (e.g., atomization energies, electron affinities, ionization potentials, proton affinities, self-interaction error (SIE) related problems, barrier heights), various reaction energies (e.g., isomerizations, Diels-Alder reactions, ozonolyses, reactions involving alkaline metals), and noncovalent interactions (water clusters, relative energies between conformers, and intermolecular and intramolecular interactions). Reference values for all subsets are based on highly accurate theoretical or experimental data, for details see the original reference.⁴⁴

As handling the large number of statistical values for such a database can be impractical, we defined a so-called weighted total mean absolute deviation (WTMAD), which combines all 30 mean absolute deviations (MADs) to one final number. For every subset the size and “difficulty” is taken into account by a factor with which each MAD is scaled. Finally, the average is taken for these scaled

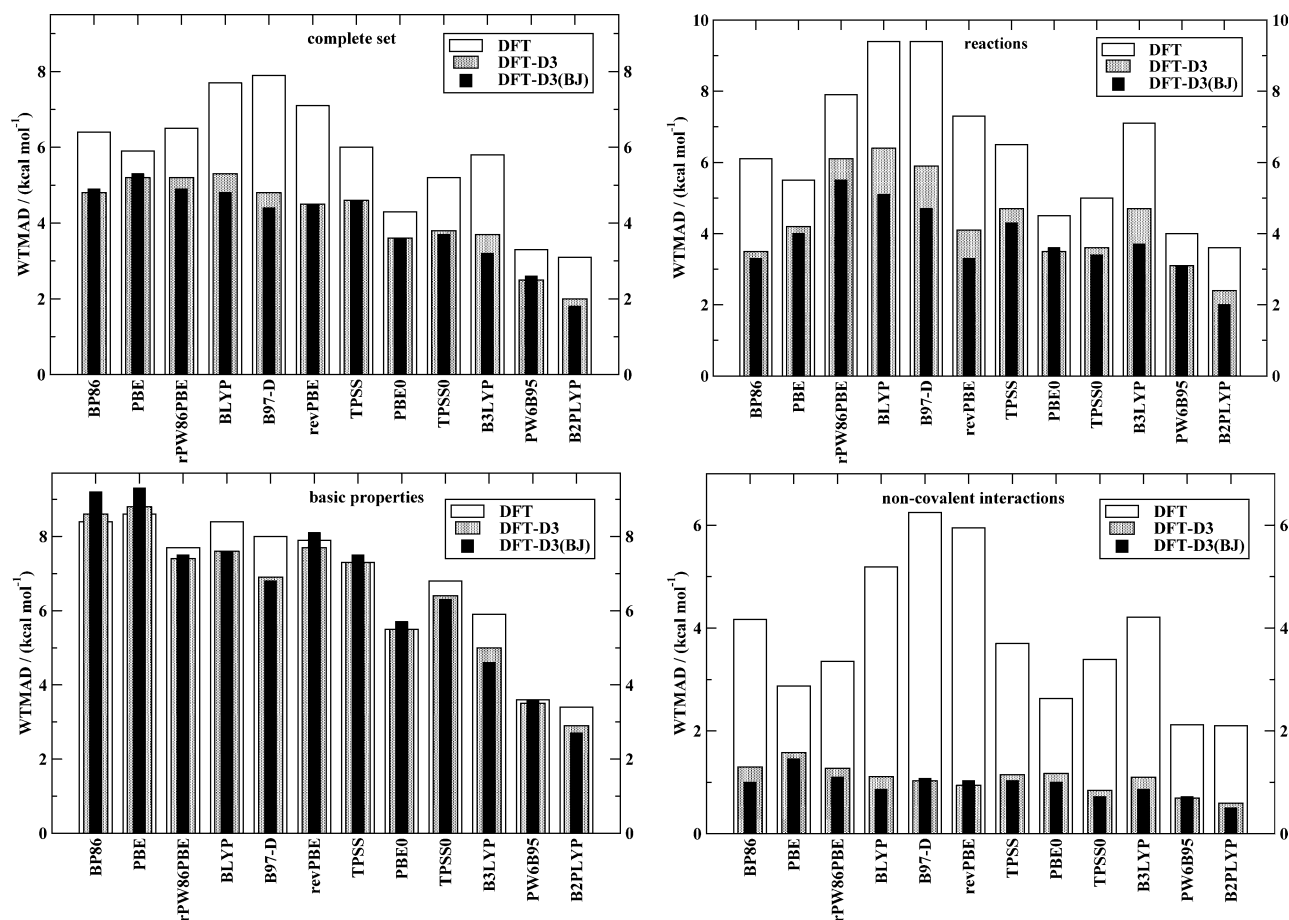


Figure 4. Weighted total mean absolute deviation (WTMAD) of 12 density functionals for (parts of) the GMTKN30 benchmark. From top to bottom: full set, basic properties of smaller molecules, reaction energies of medium and large molecules, noncovalent interactions and conformational energies.

MADs. In this work, we will adopt this idea and calculate WTMADs specifically for each of the three major sections of GMTKN30.

The results for the 12 density functionals are shown graphically for DFT-D3 and DFT-D3(BJ) in Figure 4. The errors for individual subsets when averaged over the DFs are given in Table 3.

In Figure 4, we included for comparison also plain DFT values. While it is clear that the effect of the dispersion correction is very large for noncovalent interactions (bottom of Fig. 4), it is striking that the dispersion energy is also important for “normal” thermochemistry. For the reaction part of the database (second from below in Fig. 4), both dispersion corrections improve the results significantly for all DFs. DFT-D3(BJ) is clearly better than DFT-D3 in particular for BLYP, B97-D, revPBE, and B3LYP. For the basic properties, which mainly include only small to medium-sized systems and fragmentation reactions (e.g., atomization), DFT-D3(BJ) is on average worse than DFT-D3 and in five out of 12 cases (BP86, PBE, revPBE, TPSS, PBE0) it is even slightly worse than uncorrected DFT. Although this can be expected from the basic construction of the BJ-damping when applied to standard DFs, the effect is small and for seven DFs even an improvement compared to DFT is observed. Thus, one of the important conclusions from this benchmark is that BJ-damping seems to provide

a physically correct short-range behavior of the dispersion energy even with unmodified standard correlation functionals. Even if double-counting effects of electron correlation are present in DFT-D3(BJ) (and slightly more pronounced than for zero-damping), they seem to cancel out effectively for many chemically relevant processes.

This conclusion is further supported when the results for individual subsets are analyzed. For convenience this is done as before by averaging over all considered DFs. These (total) MAD values for DFT-D3 and DFT-D3(BJ) are given in Table 3. Note that most of the noncovalent energy benchmarks in the GMTKN30 database have been used for fitting and are already discussed in the previous section. For some of the subsets (in particular the first eight) DFT-D3(BJ) is slightly worse than DFT-D3. An example are the atomization energies in W4-08 for which the total MAD increases on average by 0.21 kcal/mol (3%). This is as expected because the BJ-damping adds a correlation energy for each directly bonded atom pair, which is not present in the free atoms. If the underlying DF was designed/fit to describe atomization correctly, BJ-damping will lead to overbinding. Note, however, that there are also prominent cases of DFs, which systematically underbind molecules compared to the atoms. An example is B3LYP for which the MAD(W4-08)

Table 3. Total Mean Absolute Deviation (in kcal/mol) Averaged Over 12 Density Functionals for Subsets of the GMTKN30 Benchmark. Only Subsets That are Not Already in Table 1 are Given.

Subset	DFT-D3	DFT-D3 (BJ)
MB08-165	7.97	7.89
W4-08	6.46	6.67
G21IP	3.70	3.70
G21EA	2.99	2.98
PA	3.10	3.27
SIE11	9.86	9.90
BHPERI	2.66	2.82
BH76	6.64	6.77
BH76RC	2.91	2.93
RSE43	2.32	2.48
O3ADD6	5.11	4.98
G2RC	4.84	4.96
AL2X	3.26	2.85
NBPRC	2.68	2.34
ISO34	1.81	1.77
ISOL22	6.34	5.89
DC9	10.83	10.30
DARC	7.60	6.43
ALK6	3.71	4.23
BSR36	5.04	3.55
IDISP	4.80	3.50
WATER27	4.43	3.70
HEAVY28	0.18	0.30

value decreases from 4.19 kcal/mol with DFT-D3 to 3.65 kcal/mol for DFT-D3(BJ).

Noteable improvements by DFT-D3(BJ) are observed for reaction energy benchmarks involving relevant medium-range correlation effects. These are in particular AL2X, NBPRC, ISOL22, DARC, and BSR36 for which DFT-D3(BJ) is substantially better than DFT-D3. For example the Diels-Alder reaction energies in DARC are extremely sensitive to the description of medium-range correlation. Here, for 10 out of 12 DFs the MAD values are improved significantly by BJ-damping. The dispersion effects on noncovalent interactions in IDISP and WATER27 are also rather short-ranged and here DFT-D3(BJ) is better than DFT-D3. This is of particular importance for DFT simulations of systems involving water or H₂O clusters which are accurate even at the GGA level (e.g., with BLYP-D3(BJ)).

If the complete GMTKN30 database is considered as a good measure of performance, only for BP86, PBE and PW6B95 the BJ-damping scheme can not be fully recommended.

Structures of Noncovalently Bound Complexes

The structures of complexes in the S22 set have been optimized by Hobza and co-workers at the MP2 level using triple- ζ basis sets.³⁰ These reference data are used here to assess the quality of the two dispersion corrections for the computation of vdW and hydrogen-bonded complex geometries. We tested the functionals BLYP and TPSS as examples and employ TZVPP basis sets in the DFT calculations. The results are shown graphically in Figure 5 where the difference of the center-of-mass inter-fragment distance

ΔR_{CMA} between DFT-D3/DFT-D3(BJ) and MP2 is plotted for all 22 complexes. Note that MP2 energetically overbinds the dispersion-dominated complexes 9–15 compared to CCSD(T).³⁰ For these cases one can expect that the reference R_{CMA} values are too small, which consequently must be considered when the quality of the dispersion corrections is evaluated. For complex 14 (indole-benzene stacked) both TPSS calculations converge to a different (T-shaped with a NH- π interaction instead of π -stacked) structure than the reference, which explains the large deviations of about -1.1 Å. However, in spite of the well-known overbinding tendency of MP2 for π -stacked structures this result is understandable.

For the majority of complexes, both dispersion corrections yield very similar results. The differences between the two DFs tested are often larger. For the hydrogen-bonded systems 1–7, BJ-damping yields slightly shorter distances and the effect is a bit more pronounced for BLYP than for TPSS. BLYP-D3(BJ) is closest to the reference values but the deviations are small and within the errors of the MP2 reference data. The situation is different for the dispersion-dominated complexes 8–15 for which both DFT-D3 schemes seem to perform better than MP2 as mentioned above. Again, the differences between BLYP and TPSS (which yields slightly shorter distances) are larger than the effect of changing the damping function.

These conclusions are underlined by the comparison of DFT-D3 and DFT-D3(BJ) computed potential energy curves with accurate reference data (see Fig. 6) for three exemplary noncovalent complexes (krypton and formic acid dimers, benzene-methane). Almost two orders of magnitude of interaction energies as well as very different equilibrium distances are considered. Evidently, both damping schemes perform very similar for the krypton dimer and the benzene-methane complex. BJ-damping seems to be slightly better with the two tested functionals for the more strongly bound formic acid dimer. Here a larger effect of the damping is found at smaller distances but similar to the other complexes, the influence of the DF is more pronounced. In this regard TPSS(BLYP) is

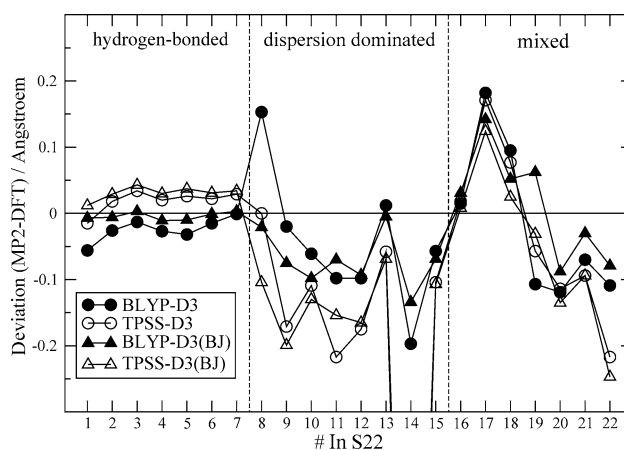


Figure 5. Difference of the center-of-mass inter-fragment distances ΔR_{CMA} between MP2 and DFT-D3/DFT-D3(BJ) for the 22 complexes in the S22 set. The values for complex 14 with TPSS are about -1.1 and do not fit the scale. Negative values indicate an overestimation of the distance by DFT-D.

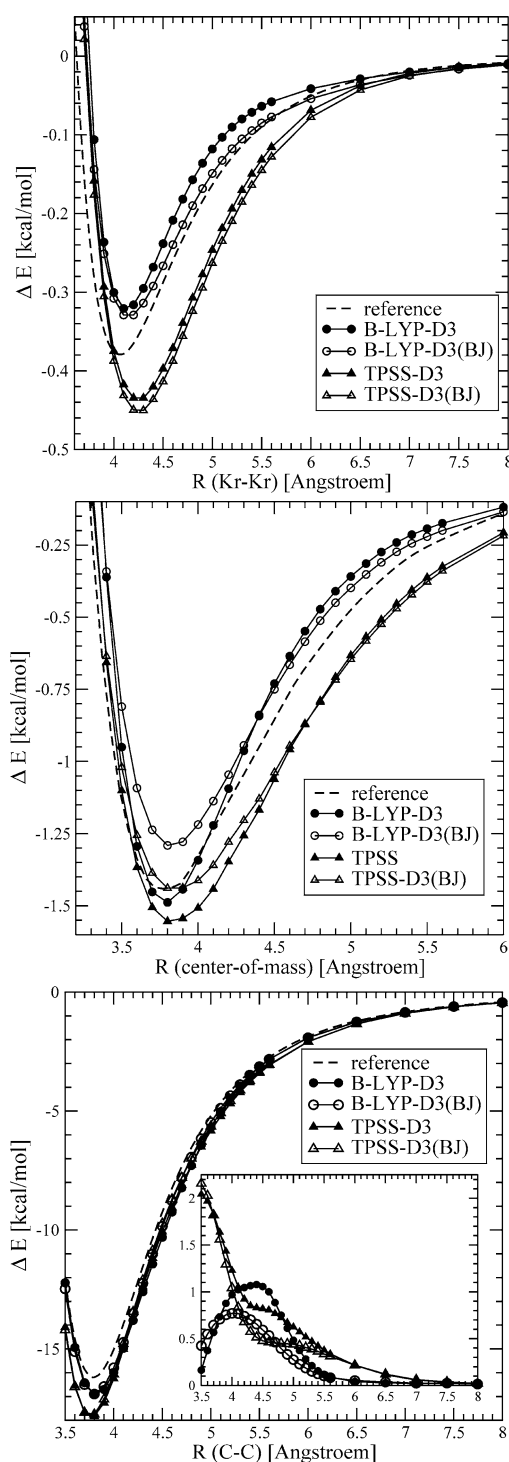


Figure 6. Interaction energy (ΔE) curves for three noncovalently bound complexes with dispersion-corrected TPSS and BLYP functionals. Accurate reference data for Kr_2 (top) and benzene-methane (C_{3v} , middle) are taken from refs. 63,64. The reference energies for the formic acid dimer (bottom) are based on valence MP2/CBS extrapolations using aug-cc-pVXZ ($X = 4, 5$) basis sets and employing MP2/aug-cc-pVTZ optimized structures. The inset shows the difference between MP2 and DFT.

over(under)binding for Kr_2 and benzene-methane while both DFs overbind the formic acid dimer (although less with DFT-D3(BJ) than with DFT-D3).

Structures of Molecules with Strong Intramolecular Dispersion Effects

As already noted, the original zero-damping in DFT-D3 can increase the distance between atoms notably when their separation is smaller than the cut-off distance R_0^{AB} but not so small that the dispersion correction is essentially zero. We investigate here four typical cases with short non-bonded contacts of various type (CC, SiN, SS, BrBr). The structures are shown in Figure 2 and the results for four representative density functionals (BLYP,^{45,46} TPSS,⁵² B3LYP,^{53,54} and PW6B95⁵⁷) in comparison to experimental data are given in Table 4.

For C_2Br_6 all un-corrected DF overestimate the BrBr distance. Without any dispersion correction, the distances are also for the other systems always in the order $\text{BLYP} > \text{B3LYP} > \text{TPSS} > \text{PW6B95}$. By applying the D3-scheme, the BrBr distance increases by 0.010–0.021 Å while it decreases by 0.003–0.017 Å for BJ-damping. All DFT-D3(BJ) results are closer to the experimental value than without dispersion correction and PW6B95-D3(BJ) almost exactly yields it.

The correct computation of the inter-ring distance in [2.2]paracyclophane is known for some time as a challenging problem.⁶⁶ Again all un-corrected DFs overestimate the CC distance, which is experimentally found to be significantly shorter than a typical CC van der Waals contact (about 3.4 Å). In this case (because the CC cut-off radius is smaller than the typical DF result of 3.1–3.2 Å) also DFT-D3 leads to a shortening. The effect, however, is too small for BLYP, B3LYP, and TPSS. DFT-D3(BJ) improves on DFT-D3 and these results are closest to the experimental value as for C_2Br_6 . The PW6B95-D3(BJ) approach performs best compared to experiment.

Table 4. Comparison of Computed and Experimental Nonbonded Distances in the Four Test Systems (for structures see Fig. 2). Values Without (DFT) and With the Two Versions of the Damping Function (DFT-D3 and DFT-D3(BJ)) are Compared. All Values are Computed using the def2-QZVP AO Basis Set.

Distance/molecule	Functional	DFT	DFT-D3	DFT-D3(BJ)	Exptl
R(BrBr) C_2Br_6	BLYP	3.531	3.556	3.506	3.42 ⁶⁵
	TPSS	3.461	3.477	3.450	
	B3LYP	3.490	3.511	3.470	
	PW6B95	3.426	3.436	3.420	
R(CC) 22PC	BLYP	3.191	3.147	3.114	3.097 ⁶⁶
	TPSS	3.127	3.103	3.103	
	B3LYP	3.157	3.126	3.093	
	PW6B95	3.097	3.084	3.086	
R(SS) S_8^{2+}	BLYP	3.499	3.563	3.222	2.86 ⁶⁷
	TPSS	3.076	3.131	3.027	
	B3LYP	3.105	3.182	2.986	
	PW6B95	2.860	2.888	2.850	
R(SiN) DTFS	BLYP	2.499	2.440	2.421	2.273 ⁶⁸
	TPSS	2.163	2.139	2.112	
	B3LYP	2.467	2.428	2.404	
	PW6B95	2.148	2.121	2.115	

This also holds for the important dihedral angle C(ar)-C-C-C(ar) which describes the (chiral) distortion from a D_{2h} to a twisted D_2 structure. This angle is small ($<5^\circ$) with dispersion-corrected BLYP, B3LYP, and TPSS functionals but about 17–19 degrees for PW6B95. The experimental value is about 13 degrees.⁶⁶ In passing we also note for this geometrical variable a positive effect of both dispersion corrections in BLYP and B3LYP calculations (increase of the angle by about 2–3 degrees).

The problem of DFT for the computed structure of S_8^{2+} has been first reported by Cameron et al. in 2000.⁶⁷ In this example, the differences between the functionals as well as the effect of the dispersion correction are by far largest. The SS distance is overestimated by 0.54 and 0.25 Å with BLYP and B3LYP, respectively. With the exception of PW6B95, the zero-damping correction leads to a significant increase of the SS distance by 0.08–0.16 Å thereby worsening the agreement to the experimental value. Switching on DFT-D3(BJ) on the other hand notably improves the pure DFT results (by 0.014–0.14 Å). For PW6B95 (which already without any correction is spot on the experimental value) the correction (correctly) has only a tiny effect.

The computation of the relative short SiN distance in (*N,N*-dimethylaminoxyl)trifluorosilane (DTFS) is also difficult with DFT⁶⁹ and the inherent errors of the functionals are larger here than the effect of the dispersion correction. While plain TPSS and PW6B95 overbind (by about 0.12 Å), BLYP and B3LYP yield too large distances (by about 0.17 Å). Similar to the case of 22PC, both versions of DFT-D3 lead to a shortening (except for PW6B95) of the SiN distance by 0.02–0.06 Å. The effect is slightly larger for BJ-damping.

If the deviations with respect to experiment are averaged over all four DFs and molecules we obtain MAD values of 0.14, 0.15, and 0.09 Å for DFT, DFT-D3, and DFT-D3(BJ). This indicates a significant improvement of the description of the intramolecular dispersion effect on molecular structures by using the BJ-damping.

Conclusions

A revision of the recently proposed DFT-D3 scheme to account for the London dispersion energy in standard DFT calculations is proposed. Because of a superior physical basis it is suggested to replace the hitherto employed so-called zero-damping function, that merges the classical and density based descriptions of electron correlation, by finite damping as introduced by Becke and Johnson (BJ-damping).

Although this method provides small but finite correlation energies for each spatially close-lying atom pair, (somewhat surprisingly) no significant “over-correlation” even with standard density functionals is found for many chemically relevant processes. This is for the first time thoroughly demonstrated by application to the GMTKN30 molecular energy database which consists of more than 1200 species and 800 reference values. In particular the revised DFT-D3(BJ) approach is better than DFT-D3 with “zero-damping” for thermochemical problems that are sensitive to medium-range correlation effects. For typical non-covalent interactions, DFT-D3(BJ) is on average for 12 standard density functionals also slightly better than DFT-D3. The advantage of the BJ-damping scheme is that

no pair-specific cut-off radii are required and that no artificial repulsive interatomic forces at short distances are present. However, after careful testing it is concluded that compared to DFT-D3 definitely one additional (global) fitting parameter is necessary. The theoretical expectations about the performance of BJ-damping regarding the description of molecular structures are mostly fulfilled. With BJ-damping better results for non-bonded distances and more clear effects of intramolecular dispersion in four representative molecules are found. Striking improvements are observed for one of the SS distances in S_8^{2+} , which is incorrectly elongated by DFT-D3 compared to pure DFT. For most of the optimized structures in the S22 set of complexes, both damping schemes lead to very similar intermolecular distances that are within the accuracy of the available reference data of only MP2 quality. In summary, DFT-D3(BJ) seems to be a small but consistent and theoretically satisfying revision of DFT-D3. Although presumably with one more free fitting parameter in the zero-damping approach the same accuracy for energies could be obtained, there still remains the advantage of the physically more sound behavior of the BJ-scheme at short distances (i.e., vanishing forces). Therefore, from now on, it will be our default in all DFT-D treatments.

Although even the BJ-damping scheme is basically empirical, this work has clearly shown that the specific choice of the mathematical form of the damping function has only a minor effect on the results if global parameters are properly adjusted. This underlines the basic physical significance of DFT-D. In practice this also means that choosing a particular exchange-correlation functional is the accuracy-determining factor. It remains to be seen if modern density-based descriptions of the dispersion energy^{70–74} can reach similar high accuracy not only for intermolecular noncovalent interactions (for which they work very well) but also for a wide range of intramolecular and thermochemical problems as both DFT-D3 variants.

Acknowledgments

This work was inspired by a CECAM workshop on vdW interactions in DFT held in Lausanne, June 2010.

References

1. Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, 1997.
2. Kaplan, I. G. *Intermolecular Interactions*; Wiley: Chichester, 2006.
3. Grimme, S. Density functional theory with London dispersion corrections, in *Wiley Interdisciplinary Reviews: Computational Molecular Science (WIREs:CMS)*; Wiley: Hoboken, New Jersey, 2010.
4. Grimme, S.; Antony, J.; Schwabe, T.; Mück-Lichtenfeld, C. *Org Biomol Chem* 2007, 5, 741.
5. Grimme, S.; Schwabe, T. *Acc Chem Res* 2008, 41, 569.
6. Huenerbein, R.; Schirmer, B.; Moellmann, J.; Grimme, S. *Phys Chem Chem Phys* 2010, 12, 6940.
7. Schwabe, T.; Huenerbein, R.; Grimme, S. *Synlett* 2010, 10, 1431.
8. Bucko, T.; Hafner, J.; Lebegue, S.; Angyan, J. G. *J Phys Chem A* 2010, 114, 11814.
9. Hepburn, J.; Scoles, G.; Penco, R. *Chem Phys Lett* 1975, 36, 451.
10. Ahlrichs, R.; Penco, R.; Scoles, G. *Chem Phys* 1977, 19, 119.
11. Cohen, J. S.; Pack, R. T. *J Chem Phys* 1974, 61, 2372.

12. Gianturco, F. A.; Paesani, F.; Laranjeira, M. F.; Vassilenko, V.; Cunha, M. A. *J Chem Phys* 1999, 110, 7832.
13. Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J Chem Phys* 2001, 114, 5149.
14. Wu, Q.; Yang, W. *J Chem Phys* 2002, 116, 515.
15. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J Chem Phys* 2010, 132, 154104.
16. von Lilienfeld, O. A.; Tkatchenko, A. *J Chem Phys* 2010, 132, 234109.
17. Grimme, S. *J Comput Chem* 2004, 25, 1463.
18. Liu, Y.; Goddard, III, W. A. *Mat Trans* 2009, 50, 1664.
19. Chai, J.-D.; Head-Gordon, M. *Phys Chem Chem Phys* 2000, 10, 6615.
20. Grimme, S. *J Comput Chem* 2006, 27, 1787.
21. Jurecka, P.; Cerny, J.; Hobza, P.; Salahub, D. R. *J Comput Chem* 2007, 28, 555.
22. Tang, K. T.; Toennies, J. P. *J Chem Phys* 1984, 80, 3726.
23. Schultz, N. E.; Zhao, Y.; Truhlar, D. G. *J Phys Chem A* 2005, 109, 11127.
24. Becke, A. D.; Johnson, E. R. *J Chem Phys* 2005, 122, 154101.
25. Johnson, E. R.; Becke, A. D. *J Chem Phys* 2005, 123, 024101.
26. Johnson, E. R.; Becke, A. D. *J Chem Phys* 2006, 124, 174104.
27. Koide, A. *J Phys B* 1976, 9, 3173.
28. Becke, A. D.; Johnson, E. R. *J Chem Phys* 2007, 127, 124108.
29. Kannemann, F. O.; Becke, A. D. *J Chem Theory Comput* 2010, 6, 1081.
30. Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys Chem Chem Phys* 2006, 8, 1985.
31. Tkatchenko, A.; Scheffler, M. *Phys Rev Lett* 2009, 102, 073005.
32. AK Grimme Research Web Site, Universität Münster, <http://www.uni-muenster.de/Chemie.oc/grimme/>. Accessed February 1, 2011.
33. TURBOMOLE, version 5.9: R. Ahlrichs et al., Universität Karlsruhe 2009. See <http://www.turbomole.com>. Accessed February 1, 2011.
34. Weigend, F.; Furche, F.; Ahlrichs, R. *J Chem Phys* 2003, 119, 12753.
35. Weigend, F.; Ahlrichs, R. *Phys Chem Chem Phys* 2005, 7, 3297.
36. Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. *J Chem Phys* 2003, 119, 11113.
37. Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chem Phys Lett* 1993, 213, 514.
38. Weigend, F. *Phys Chem Chem Phys* 2002, 4, 4285.
39. Weigend, F.; Häser, M. *Theor Chem Acc* 1997, 97, 331.
40. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor Chem Acc* 1997, 97, 119.
41. Weigend, F.; Köhn, A.; Hättig, C. *J Chem Phys* 2002, 116, 3175.
42. Weigend, F. *Phys Chem Chem Phys* 2006, 8, 1057.
43. Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D. *J Chem Phys* 2010, 132, 144104.
44. Goerigk, L.; Grimme, S. *J Chem Theor Comput*. DOI: 10.1021/ct100466k.
45. Becke, A. D. *Phys Rev A* 1988, 38, 3098.
46. Lee, C.; Yang, W.; Parr, R. G. *Phys Rev B* 1988, 37, 785.
47. Perdew, J. P. *Phys Rev B* 1986, 33, 8822.
48. Perdew, J. P. *Phys Rev B* 1986, 34, 7406.
49. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys Rev Lett* 1996, 77, 3865.
50. Zhang, Y.; Yang, W. *Phys Rev Lett* 1998, 80, 890.
51. Murray, E. D.; Lee, K.; Langreth, D. C. *J Chem Theory Comput* 2009, 5, 2754.
52. Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys Rev Lett* 2003, 91, 146401.
53. Becke, A. D. *J Chem Phys* 1993, 98, 5648.
54. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J Phys Chem* 1994, 98, 11623.
55. Adamo, C.; Barone, V. *J Chem Phys* 1999, 110, 6158.
56. Grimme, S. *J Phys Chem A* 2005, 109, 3067.
57. Zhao, Y.; Truhlar, D. G. *J Phys Chem A* 2005, 109, 5656.
58. Grimme, S. *J Chem Phys* 2006, 124, 034108.
59. Kannemann, F. O.; Becke, A. D. *J Chem Theory Comput* 2009, 5, 719.
60. Krieg, H.; Grimme, S. *Mol Phys* 2010, 108, 2655.
61. Grimme, S. *Org Lett* 2010, 12, 4670.
62. Goerigk, L.; Grimme, S. *J Chem Theor Comput* 2010, 6, 107.
63. Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. *J Phys Chem A* 2009, 113, 10146.
64. Slavicek, P.; Kalus, R.; Paska, P.; Odvarkova, I.; Hobza, P.; Malijevsky, A. *J Chem Phys* 2003, 119, 2102.
65. Mandel, G.; Donohue, J. *Acta Cryst* 1972, B28, 1313.
66. Grimme, S. *Chem Eur J* 2004, 10, 3423.
67. Cameron, T.; Deeth, R.; Dionne, I.; Du, H.; Jenkins, H.; Krossing, I.; Passmore, J.; Roobottom, H. *Inorg Chem* 2000, 39, 5614.
68. Mitzel, N.; Losehand, U.; Wu, A.; Cremer, D.; Rankin, D. *J Am Chem Soc* 2000, 122, 4471.
69. Gerenkamp, M.; Grimme, S. *Chem Phys Lett* 2004, 392, 229.
70. Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D.; Lundqvist, B. *Phys Rev Lett* 2004, 92, 246401.
71. Langreth, D. *et al.* *J Phys Condens Matter* 2009, 21, 084203.
72. Vydrov, O. A.; Vorhis, T. V. *Phys Rev Lett* 2009, 103, 063004.
73. Vydrov, O. A.; Voorhis, T. V. *Phys Rev A* 2010, 81, 062708.
74. Klimeš, J.; Bowler, D.; Michaelides, A. *J Phys Cond Mat* 2010, 22, 022201.