JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

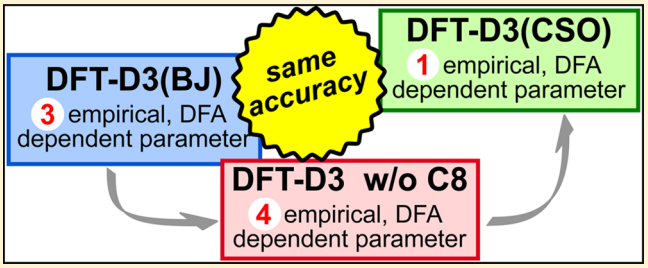
Reformulation of the D3(Becke−Johnson) Dispersion Correction without Resorting to Higher than C_6 Dispersion Coefficients

Heiner Schröder, Anne Creon, and Tobias Schwabe*

Center for Bioinformatics and Institute of Physical Chemistry, University of Hamburg, Bundesstraße 43, 20146 Hamburg, Germany

Supporting Information

ABSTRACT: A reformulated version of Grimme's most recent DFT dispersion correction with Becke−Johnson damping (DFT-D3(BJ)) is presented, which only depends on C_6 dispersion coefficients. The role of the higher order correction terms in the DFT-D3(BJ) model is critically investigated, and a sigmoidal interpolation function for adjusting to different density functional approximations (DFA) is employed alternatively, while keeping finite damping of Becke and Johnson. For the proposed C_6-only dispersion correction scheme, only one parameter needs to be fitted per DFA (instead of three for DFT-D3(BJ)). Eight standard DFAs from different classes are parametrized and evaluated. In comparison to DFT-D3(BJ), one of the most accurate corrections up to date, the new correction shows only negligible deviations in accuracy for the huge GMTKN30 benchmark set.



1. INTRODUCTION

Kohn−Sham density functional theory has become the most popular electronic structure method in computational chemistry. One of the most important reasons for its success is the advantageous scaling behavior compared to wave function based methods, which allows for the calculation of large systems up to 10^3 atoms for practical applications such as biological systems or nanoscaled compounds. In such systems, dispersion interactions often play a significant role. Unfortunately, the proper description of London dispersion interactions is a major weakness of standard density functional approximations (DFAs).[1] Standard DFAs fail to reproduce the correct interaction potentials, especially the R^{-6}-asymptotics. When dispersion is the dominant interaction, complexes are usually bound too weakly,but may in some cases even be purely repulsive, while hydrogen bonds are fairly well described.[2]

To make up for this, a variety of different dispersion correction schemes has been proposed. Analogue to the "Jacob's ladder"[3] of DFAs, Klimeš and Michaelides proposed a "stairway to heaven"[4] of dispersion correction schemes. Each step of the stairway introduces more robust, but computationally more demanding, correction schemes. For an overview over different dispersion correction schemes see refs 4 and 5. This work is focused on the simple yet accurate atom pairwise empirical dispersion corrections, often called DFT-D, and especially on the DFT-D3(BJ) scheme.

The simplicity of DFT-D like corrections stems mostly from the fact that the dispersion energy E_{disp} is computed independently from the Kohn−Sham energy E_{KS} to yield the total energy E_{tot} of the system such that there is no direct influence on the electron density.

$$E_{tot} = E_{KS} + E_{disp} \tag{1}$$

For DFT-D like schemes, the dispersion energy is treated as pairwise additive, while three-body terms are usually, but not always,[6,7] neglected. The energy expression has the following general form:

$$E_{disp} = -\sum_{AB} \sum_{n=6,8,10,\ldots} \frac{C_n^{AB}}{R_{AB}^n} f_{damp,n}(R_{AB}, A, B) \tag{2}$$

The success of such an empirical correction is based on three major parts which are the subject of its principal design. First, it has to restore the correct physical asymptotics through applying C_nR^{-n}-potentials which comes from the perturbation theory analysis of the interaction of two independent charge clouds.[8] Here, the key ingredients are the C_n^{AB} dispersion coefficients of order n and are specific for each atom pair AB. In principal, the series goes to infinite order and therefore has to be truncated at an appropriate level. The individual contributions from each atom pair have to be summed up to obtain the total dispersion energy. Second, the form of the potentials is only correct for sufficiently large atom−atom distances R_{AB} but diverges for small distances. To avoid this, a damping function $f_{damp,n}$ has to be applied whose actual form is a matter of choice. Third, to be combinable with general DFAs, a way to adapt the actual dispersion correction for a given DFA has to be included. This is normally realized with DFA dependent parameters which have to be fitted empirically. Only a well balanced combination of these parts leads to a reliable correction, and several different

Received: April 29, 2015
Published: May 29, 2015

DOI: 10.1021/acs.jctc.5b00400
J. Chem. Theory Comput. 2015, 11, 3163−3170

versions have been proposed. For an overview, see again refs 4 and 5.

For simplicity, many DFT-D schemes include only C_6 terms,[9−12] but also faster decaying terms (C_8R^{-8}, $C_{10}R^{-10}$) have been included in earlier versions[13,14] and see a revival in recent correction schemes.[6,15−19] Grimme concluded in a recent review about DFT dispersion correction that "some kind of consensus has been reached in that C_6 alone is not sufficient to describe medium/short-range dispersion".[5]

Although the C_6 (and higher order) coefficients are a fundamental ingredient of DFT-D like corrections, there is no unique way to obtain these. Different methods of varying complexity have been proposed.[6,12,17−19] One difficulty arises from the fact that the exact computation of coefficients with response (or time-dependent) approaches can only be carried out for the whole system (i.e. atoms, molecules, complexes, and so on) but for DFT-D like corrections the coefficients have to be distributed among the atoms constituting the system. There is in general no unique way for such a distribution. Further, when derived from perturbation analysis, non- or weakly interacting systems are assumed. It is not guaranteed that coefficients obtained for e.g. free atoms can easily be transferred to bound states in a molecule.[6] Note that the validity of the atom pairwise additivity has been challenged recently—at least for (metallic) nanonstructures.[20] These problems have been overcome to some extent in recent work where dispersion coefficients for whole systems even to higher orders are accurately computed in an efficient way.[21] However, the underlying model has not been generalized to arbitrary systems, and, at the moment, it is limited to spherical (but possibly hollow) structures as well as to intermolecular but not intramolecular interactions. Nevertheless, promising results for nanoscopic systems have been presented.[22]

For the damping function, different approaches have been chosen. Older DFT-D schemes apply damping functions which go to zero for short distances. The individual form varies from the more complex Tang−Toennies variant,[23] to the Fermi-form introduced by Wu and Yang[24] or the Chai−Head-Gordon power function[25] or a very flexible form, which covers most of these cases.[26] More recent developments take Koide's analysis into account which shows that the dispersion energy should become constant for short interatomic distances.[27] Becke and Johnson introduced the finite damping first in their model,[17] and it was also adopted for the DFT-D3 scheme.[16]

Besides ensuring the correct physics for the short- and long-range asymptotics, the damping function also needs a certain flexibility to allow for the different behavior of individual DFAs—especially, if DFAs themselves should not be reparametrized for the use with a dispersion correction. Normally, three different regions are identified qualitatively: the short-range region where DFAs should already give correct interaction energies (and there should be no contribution from a dispersion correction), the long-range region, where DFAs are expected to fail completely and therefore it should be described completely by the dispersion correction, and the midrange region, where DFAs still contribute to the interaction energy but have to be supported to some extent by the empirical correction.

Because of this important function for the correction scheme, the actual performance of a given combination of the DFT-D scheme and DFA can be quite sensitive to the choice of damping function and the parameters therein[28] and therefore adjusting the parameters needs careful fitting. While the adjustment of the damping function is one way to combine a DFT-D scheme with a given DFA,[11,12,18] further global scaling parameters have also been introduced for that purpose.[9,10] In the end, for some DFT-D schemes it was found that a combination of both approaches leads to the most successful correction schemes and is therefore used.[6,11,16]

As has been pointed out before, the actual design choices of a DFT-D correction scheme are not independent. It has been noted that the actual form of a damping function can model the effect of higher order correction terms (and vice versa)[4,11] and also that the actual fitting of a correction scheme can compensate for the performance dependence on the damping function.[6] In this paper, the interdependence of the ingredients shall be elucidated by a careful analysis of the very popular DFT-D3(BJ) scheme.[16] It combines many of the latest suggestions of how a DFT-D scheme should be designed: it accounts not only for C_6 but also for C_8 terms, the atomic dispersion coefficients are system dependent (though not computed directly from the electron density), and it employs a finite damping scheme. It also has been successfully combined with a broad variety of DFAs. Since in the DFT-D3(BJ) scheme, the higher order dispersion coefficients are recursively calculated from the C_6 coefficient[29,30] in a linear fashion, inclusion of those terms does not seem to be necessary in principal. In this work the role of the higher order contributions in the DFT-D3(BJ) scheme is investigated. The empirical nature of the C_8 term for this specific scheme and how it is related to adjusting the correction to a given DFA is emphasized. In addition, an alternative correction scheme is introduced, which only depends on C_6 coefficients alone and also reduces the number of empirical parameters in comparison to DFT-D3(BJ) to further reduce the empiricism of the correction.

2. THEORY

The DFT-D3(BJ) energy is given by

$$E^{D3(BJ)} = -\sum_{AB} s_6 \frac{C_6^{AB}}{R_{AB}^6 + (a_1 R_0^{AB} + a_2)^6} + s_8 \frac{C_8^{AB}}{R_{AB}^8 + (a_1 R_0^{AB} + a_2)^8} \tag{3}$$

where the scaling factor $s_6 = 1$ for standard DFAs to ensure the right long-range asymptotics where only the C_6-term is relevant. a_x and s_8 are functional dependent parameters fitted to noncovalent interaction energies. R_0^{AB} is an empirical but fixed cutoff radius for atom pair AB.

To investigate the role of the C_8^{AB} term, we emphasize the following simple relation used in D3(BJ)

$$C_8^{AB} = 3 k_{AB} C_6^{AB} \tag{4}$$

where k_{AB} is an atom pair dependent coefficient and

$$k_{AB} = s_{42} \sqrt{\sqrt{Z_A}\left(\frac{\langle r^4\rangle^A}{\langle r^2\rangle^A}\right)\sqrt{Z_B}\left(\frac{\langle r^4\rangle^B}{\langle r^2\rangle^B}\right)} \tag{5}$$

s_{42} is an empirical, redundant linear scaling parameter of value 1/2, which could, along with the factor 3, be merged with the empirical s_8 parameter. $\langle r^4\rangle^A$ and $\langle r^2\rangle^A$ are quadrupole and dipole moment expectation values, respectively, computed for the free atom A. Their ratio is scaled with a further empirical

DOI: 10.1021/acs.jctc.5b00400
J. Chem. Theory Comput. 2015, 11, 3163−3170

scaling term $\sqrt{Z_A}$ which depends on the nuclei charge Z_A of atom A. Note that the cutoff radius R_0^{AB} is defined as

$$R_0^{AB} = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}} = \sqrt{\frac{3}{2}\sqrt{\sqrt{Z_A}\left(\frac{\langle r^4\rangle^A}{\langle r^2\rangle^A}\right)\sqrt{Z_B}\left(\frac{\langle r^4\rangle^B}{\langle r^2\rangle^B}\right)}} \tag{6}$$

but, in principle, other choices to define R_0^{AB} are possible and have also been tested. All atom-dependent parameters for the DFT-D3 approach are precomputed and tabulated. For a complete discussion of the derivation, see ref 6 and ref 16 and references therein.

Plugging eq 4 into eq 3, one obtains

$$\begin{aligned} E_{disp}^{D3(BJ)} &= -\sum_{AB} s_6 \frac{C_6^{AB}}{R_{AB}^6 + (a_1 R_0^{AB} + a_2)^6} + s_8 \frac{3k_{AB}C_6^{AB}}{R_{AB}^8 + (a_1 R_0^{AB} + a_2)^8} \\ &= -\sum_{AB}\left[s_6 + \frac{3s_8 k_{AB}(R_{AB}^6 + (a_1 R_0^{AB} + a_2)^6)}{R_{AB}^8 + (a_1 R_0^{AB} + a_2)^8}\right] \\ &\quad \times \frac{C_6^{AB}}{R_{AB}^6 + (a_1 R_0^{AB} + a_2)^6} \end{aligned} \tag{7}$$

The terms within square brackets can be regarded as an interpolation function f^{inp} which is introduced by the use of the higher order correction term. It interpolates between the $R \to 0$ and $R \to \infty$-limits which are

$$\lim_{R_{AB}\to 0} f^{inp}(R_{AB}) = s_6 + \frac{3s_8 k_{AB}}{(a_1 R_0^{AB} + a_2)^2} = const \tag{8}$$

$$\lim_{R_{AB}\to \infty} f^{inp}(R_{AB}) = s_6 \tag{9}$$

Together with the damping term in the denominator of the leading C_6 term, f^{inp} ensures that the dispersion correction becomes finite for small distances and has the asymptotic form $C_6 R_{AB}^{-6}$ for large distances. Note that the same functional form of $E_{disp}^{D3(BJ)}(R_{AB})$ could have been achieved without f^{inp}, and therefore it is not necessary for the correct physics. On the other hand, the actual value of $E_{disp}^{D3(BJ)}(R_{AB})$ in the short and midrange region heavily depends on the empirical fitting parameters s_8 and a_x. Therefore, the main purpose of the higher order term in the DFT-D3(BJ) scheme is the introduction of flexibility to adjust the correction for a given DFA and to compensate for possible shortcomings because of the neglect of higher order terms. Actually, the role of the s_8 parameter has already been emphasized in the original DFT-D3 paper,[6] but to the best of the authors' knowledge the similar role of the damping parameters in the DFT-D3(BJ) scheme has not been pointed out that clear.

f^{inp} is affected the most by the empirical fitting of the correction scheme. While there is theoretical reasoning for the use of the higher order terms to correct interaction energy deficiencies, the correct mathematical form for the combination of the correction with different DFAs is not known. Of course, any chosen function should yield the same limits as found in eqs 8 and 9.

In the following, alternative choices are discussed and investigated. By inspecting f^{inp}, one finds that it could be approximated as a function of the form $(1 + 1/(R_{AB}^2 + const))$. Different corrections schemes with such a quadratic interpolation function were tested but could not reach the accuracy of DFT-D3(BJ). (Results are not shown here.)

Another typical function that may satisfy the requirements and is often applied for interpolation in general is a sigmoidal function. This form exhibited promising results in first tests. As the dispersion correction depends only on C_6 coefficients now, it has been named C-Six-only DFT-D3(CSO) (or short D3(CSO)). The DFT-D3(CSO) dispersion energy is given by

$$\begin{aligned} E_{disp}^{D3(CSO)} &= -\sum_{AB}\left[s_6 + \frac{a_1}{1 + \exp(R_{AB} - a_2 R_0^{AB})}\right] \\ &\quad \times \frac{C_6^{AB}}{R_{AB}^6 + (a_3 R_0^{AB} + a_4)^6} \end{aligned} \tag{10}$$

This introduces four empirical DFA dependent parameters. a_1 determines the deviation from the ideal s_6 value of 1 at short distances which is on the one hand related to the general capability of a given DFA to cover dispersion effects for short distances and on the other hand compensates for missing contributions from higher order terms. Too repulsive DFAs will have a large a_1, and those which capture some or most of the dispersion contribution at short distances will have a low or even negative a_1. But as long as $a_1 > -s_6$, the correction is always attractive. a_2 (along with the damping term) controls at which distance the dispersion contribution will start to deviate from the ideal R^{-6} curve. For $R_{AB} = a_2 R_0^{AB}$, the sigmoidal function has its turning point and $f^{inp} = s_6 + 0.5a_1$.

Because the functional behavior is dependent on the electron density (and therefore the overlap of (atom) densities) this should be reflected in the dispersion correction. In the given model, this is captured by introducing a distance dependence on R_0^{AB} for a given atom pair AB. The actual value of a_2 depends on the functional performance and the need to model higher order terms which become more important at shorter distances.

Finally, a_3 and a_4 scale the damping term. The term determines the final contribution for small R_{AB} and also regulates at which distance the damping becomes effective. Again, this should also be related to the (atom) density overlap, and, therefore, this term depends also on R_0^{AB}. The four parameters are not fully decoupled, and qualitatively very similar curves can be obtained with different parameter sets. Although eq 10 defines our general working model, slight changes to this form were made during the fitting procedure and will be shown and discussed later on.

It should also be noted that in principal higher order terms can be kept with the correct expansion coefficient $s_8 = 1$, and, in addition, a suitable interpolation function can be introduced for the combination of that correction with a given DFA. But as long as a linear recursion relation between C_6 and C_8 coefficients is assumed—like for the DFT-D3 method—, this might only increase the complexity of the correction without leading to higher accuracy. For other correction schemes where higher order coefficients are obtained differently such an approach might be worth investigating.

3. COMPUTATIONAL DETAILS

Geometries for the GMTKN30-[31,32] and S66 set[33] were taken from the corresponding references. All energies in the SCF steps have been converged to 10^{-7} au. All DFT calculations were performed with a locally modified version of TURBOMOLE 6.4[34] using the integration grid m4.[35] For all calculations the def2-QZVPP basis set[36] was used. For reactions of the G2EA subset augmented diffuse functions from corresponding Dunning basis sets[37] were employed to account for the diffuse electronic structure of anions. Dispersion corrections were computed with a locally modified version of

DOI: 10.1021/acs.jctc.5b00400
J. Chem. Theory Comput. 2015, 11, 3163−3170

Grimme's dftd3-program.[6,16] For the perturbation part of the double-hybrid functional the resolution-of-the-identity-approximation with the corresponding auxiliary basis sets[38,39] was used. Additionally, the frozen core approximation was applied as it is less prone to basis-set incompleteness error.[40]

The equilibrium distance of the Kr-Dimer was determined on the MP2/aug-cc-pVQZ-level using TURBOMOLE. Calculations of the interaction energies were performed with ORCA 3.02,[41] because of technical issues with TURBOMOLE. Energies in the SCF steps have been converged to 10^{-8} au.

The D3(CSO) damping parameters were optimized by a least-squares-fit to reference energies of the S66 benchmark set of noncovalent interactions[33] using a local Python implementation for the computation of the dispersion corrections.

4. RESULTS AND DISCUSSION

The function parameters need to be determined by fitting to noncovalent interaction energies. Therefore, we choose the S66 benchmark set[33] of noncovalent interactions, which consists of 66 weakly bound model complexes from biochemistry, including H-bonds, aliphatic-aliphatic-, $\pi-\pi$- and $\pi-$aliphatic-interactions in varying distances. Parameter optimization has been performed for a variety of different standard functionals from different classes. Three generalized gradient approximations (GGAs): BP86,[42−44] BLYP,[42,45]PBE,[46] one meta-GGA: TPSS,[47] three hybrid GGAs: B3LYP,[48,49] PBE0,[50] PW6B95[51] and one double-hybrid functional: B2PLYP.[52]

In a lot of cases more than one minimum on the error hypersurface could be observed for fitting to the S66 set alone. The choice of the minimum had significant effect on later benchmarking. Especially if negative values were found for a_3 and a_4, the results became inconsistent. So a first step was to square the parameters a_3 and a_4, so that positive values could be ensured. Since multiple minima still remained, two reactions of alkaline metal (Li,Na) metal clusters M_8 decomposing into the corresponding dimers, which showed strong dependence on parameter choice, had been added to the fit set and finally led to a single minimum on the error hypersurface for all functionals parametrized here.

We started the optimization procedure with the form of eq 10. The scaling parameter s_6 was set to unity for all DFAs but B2PLYP, where s_6 was lowered by the amount of WF-correlation included to attribute for its inherent nonlocal correlation part. The parameter a_3 approached zero for all functionals and was thus removed. Probably, it is sufficient when the atom pair dependence is covered in the interpolation function and the C_6 coefficients. The constant parameter a_4 was found to be 2.5 ± 0.2 for all functionals (except for B2PLYP) and could be fixed at 2.5 without worsening the overall results. There was no influence of the parameter a_2, if a_1 was properly fit. To obtain a value for a_2, we fixed a_2 at different values and optimized a_1. We choose the lowest value of a_2, which had no significantly negative influence for all functionals. Coincidentally, it was found to be 2.5 as well. The final expression for the D3(CSO) dispersion energy is given by

$$E_{disp}^{D3(CSO)} = -\sum_{AB}\left[s_6 + \frac{a_1}{1+\exp(R_{AB} - 2.5R_0^{AB})}\right]\frac{C_6^{AB}}{R_{AB}^6 + (2.5^2)^6} \quad (11)$$

Here, the parameter a_1 remained as the only DFA dependent parameter. It only influences the scaling of the sigmoidal function. The final optimized parameter values are given in Table 1.

Table 1. Optimized Parameter Values (in atomic units) for 8 DFAs

functional	a_1	functional	a_1
BLYP	1.28	B3LYP	0.86
BP86	1.01	PBE0	0.20
PBE	0.24	PW6B95	−0.15
TPSS	0.72	B2PLYP[a]	0.24

[a] $s_6 = 0.73$

The values for a_1 correspond to the "repulsiveness" of the functional as does its D3(BJ)-counterpart s_8, which also acts as a scaling factor to the interpolation function.[16]

For comparison, both dispersion corrections for the He dimer are plotted in Figure 1. As can be seen, both functional forms are qualitatively very similar. Differences are mainly due to parametrization and not due to the different functional forms.

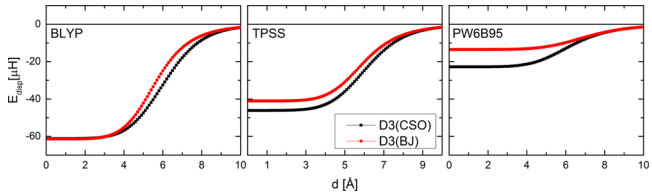


Figure 1. Dispersion correction contributions for the He dimer and three representative DFAs.

4.1. Fit Set. The final results for the fit set in comparison to uncorrected or D3(BJ) results is shown in Table 2. First of all, a significant reduction of the MD and MAD is observed for both corrections for every functional, while the error range only decreases for the cases involving the LYP correlation functional. D3(BJ) performs slightly better than D3(CSO) for most functionals, but this stems almost exclusively from the better results for the decomposition of the alkaline metal clusters. Figure 2 shows the correlation between both correction schemes for three representative functionals. Deviations can only be observed for the mentioned reactions.

4.2. Dissociation. In the former section, dispersion interactions were only looked at at equilibrium distances. To evaluate the performance of the new dispersion correction for varying distances, three representative noncovalently bound complexes have been chosen: the Krypton-dimer, an ethene-benzene-$\pi-\pi$-complex, and a double H-bonded uracil-peptide model complex. The last two were taken from the S66x8 benchmark set.[33] The interaction energy curves are shown in Figure 3.

In general, both corrections lead to qualitatively and quantitatively correct potential curves. Deviations are within the accuracy which can be expected from empirical correction schemes. While both corrections perform very similar for TPSS in all three cases, they diverge slightly for BLYP. Interestingly, whether BLYP-D3(CSO) gives larger or smaller interaction energies (relative to BLYP-D3(BJ)) is not systematic but system dependent. In any case, the new model shows the desired behavior and is not less accurate than its predecessor.

4.3. Thermodynamic Benchmark. Finally, the new D3(CSO) correction has been assessed based on the extensive

DOI: 10.1021/acs.jctc.5b00400
J. Chem. Theory Comput. 2015, 11, 3163−3170

Table 2. Statistical Error Values of All Parametrized DFAs for the Fit Set[a]

	BP	BLYP	PBE	TPSS	B3LYP	PBE0	PW6B95	B2PLYP
none								
MD	−3.6	−4.4	−1.9	−2.9	−3.6	−1.9	−1.4	−1.8
MAD	3.6	4.4	2.2	3.0	3.6	2.1	1.7	1.8
Δ_{err}	9.5	22.4	17.7	13.7	19.8	13.7	13.0	7.8
D3(BJ)								
MD	0.4	0.0	0.6	0.3	0.2	0.5	0.3	0.1
MAD	0.6	0.3	0.7	0.5	0.4	0.6	0.6	0.2
Δ_{err}	12.8	9.8	16.9	13.4	7.7	13.1	14.5	3.1
D3(CSO)								
MD	0.4	0.2	0.7	0.6	0.2	0.6	0.4	0.1
MAD	0.8	0.4	0.8	0.8	0.3	0.7	0.6	0.3
Δ_{err}	18.9	10.4	22.9	21.1	7.4	18.7	16.9	5.1

[a]Errors given are the mean deviation (MD), mean absolute deviation (MAD), and the error range Δ_{err} and are given without correction and with D3(BJ)- and D3(CSO)-correction. All values are in kcal mol^{-1}.

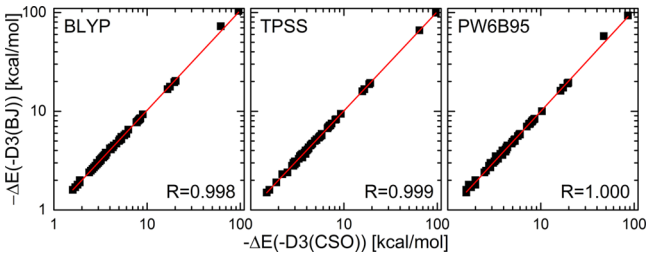


Figure 2. Correlation-plot of reaction energies obtained by D3(BJ) (*y*-axis) and D3(CSO) (*x*-axis) dispersion correction for three functionals (logarithmic scaling used).

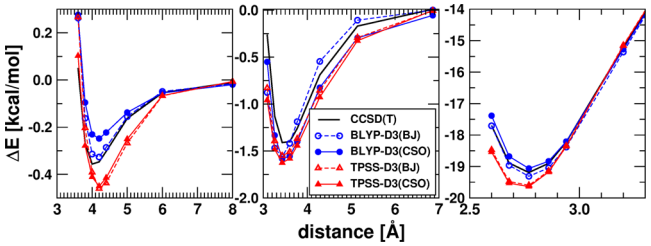


Figure 3. Interaction energy curves (ΔE) for noncovalently bound complexes computed with the TPSS and BLYP functional, both with and without dispersion correction: Kr dimer (left), ethene-benzene complex (middle), uracil-peptide complex (right). Data points are at 0.9, 0.95, 1, 1.05, 1.1, 1.25, 1.5, 2 times the equilibrium distance except for the uracil-peptide complex where only the first six data points are shown. Reference values for the Krypton dimer were calculated on the CCSD(T)/augccpVQZ level. For the other two complexes CCSD(T)/CBS values were taken from ref 33.

GMTKN30 test set (see Figure 4). For the evaluation, we also adopted the division of the whole GMTKN30 set into the same subcategories as proposed in ref 32. All computed data can be found in the Supporting Information.

As expected, the largest relative differences between uncorrected and corrected values can be found for the noncovalent interactions. Surprisingly, PBE does not benefit from any correction, whereas all other tested functionals benefit by at least 0.5 kcal mol^{-1}. The disadvantageous correction for PBE stems almost completely from the Water27-subset. PBE already overbinds the water complexes of the subset, so adding additional dispersion energy further worsens the results. Overall, both corrections perform almost equally well with the exception of the BLYP- and B3LYP-functional for which D3(CSO) improves the MAD by 0.2−0.3 kcal mol^{-1} compared to D3(BJ). The deviations are again almost only due to the Water27-subset.

Even though the description of the Water27-subset is improved for D3(CSO), this does not hold for H-bonded complexes in general as can be seen by the H-bonded complexes of the S66 set. Here, the MAD is 0.2 vs 0.3 kcal mol^{-1} for BLYP with D3(BJ) or D3(CSO), respectively, while it is 6.7 and 5.2 kcal mol^{-1} for the Water27 subset. For B3LYP, the same trends can be observed with MADs of 0.4 (D3(BJ)) and 0.3 (D3(CSO)) kcal mol^{-1} for the S66 H-bond subset and MADs of 6.7 (D3(BJ)) and 5.4 (D3(CSO)) kcal mol^{-1} for the Water27 subset. Finally, for B2PLYP almost no difference between the two corrections can be found. For both cases, the MAD is 0.2 kcal mol^{-1} for the S66 H-bond subset and 4.6 (D3(BJ)) vs 4.5 (D3(CSO)) kcal mol^{-1} for the Water27 subset.

On an absolute scale, reaction energies benefit even more from dispersion correction. The results for all functionals improve by at least 0.5 kcal mol^{-1}(PBE) and up to about 3 kcal mol^{-1} (BLYP). The difference between both correction schemes is always within the 0.1 kcal mol^{-1} range which is statistically no significant difference in this case.

Least effect is observed for the energies of the basic properties section. This is to be expected, since most energies are calculated for small to medium-sized molecules, where dispersion does not play an important role. For BP, PBE and TPSS both dispersion corrections even worsen the MAD results. The maximum increase is 0.3 kcal mol^{-1} for BP-D3(CSO) which is too small to indicate any general problems. Comparing D3(BJ) and D3(CSO), the deviations between MADs again do not exceed 0.1 kcal mol^{-1} and are thus negligible.

The overall similarity of both corrections can also be found, when looking at the errors for the individual subsets, which are shown in Figure 5 for the representative DFAs BLYP, TPSS, and PW6B95. Significant differences between the correction schemes for all three functionals can only be found for the AL2X (#3) and ALK6 (#4) subsets, though it depends on the DFA which correction scheme gives better results. Interestingly, for subset #4, dispersion correction increases the MAD for some functionals. The question, if this is a shortcoming of the dispersion correction or if an improved description of London dispersion abrogates a fortunate error compensation for the corresponding DFA cannot be unraveled here.

DOI: 10.1021/acs.jctc.5b00400
J. Chem. Theory Comput. 2015, 11, 3163−3170

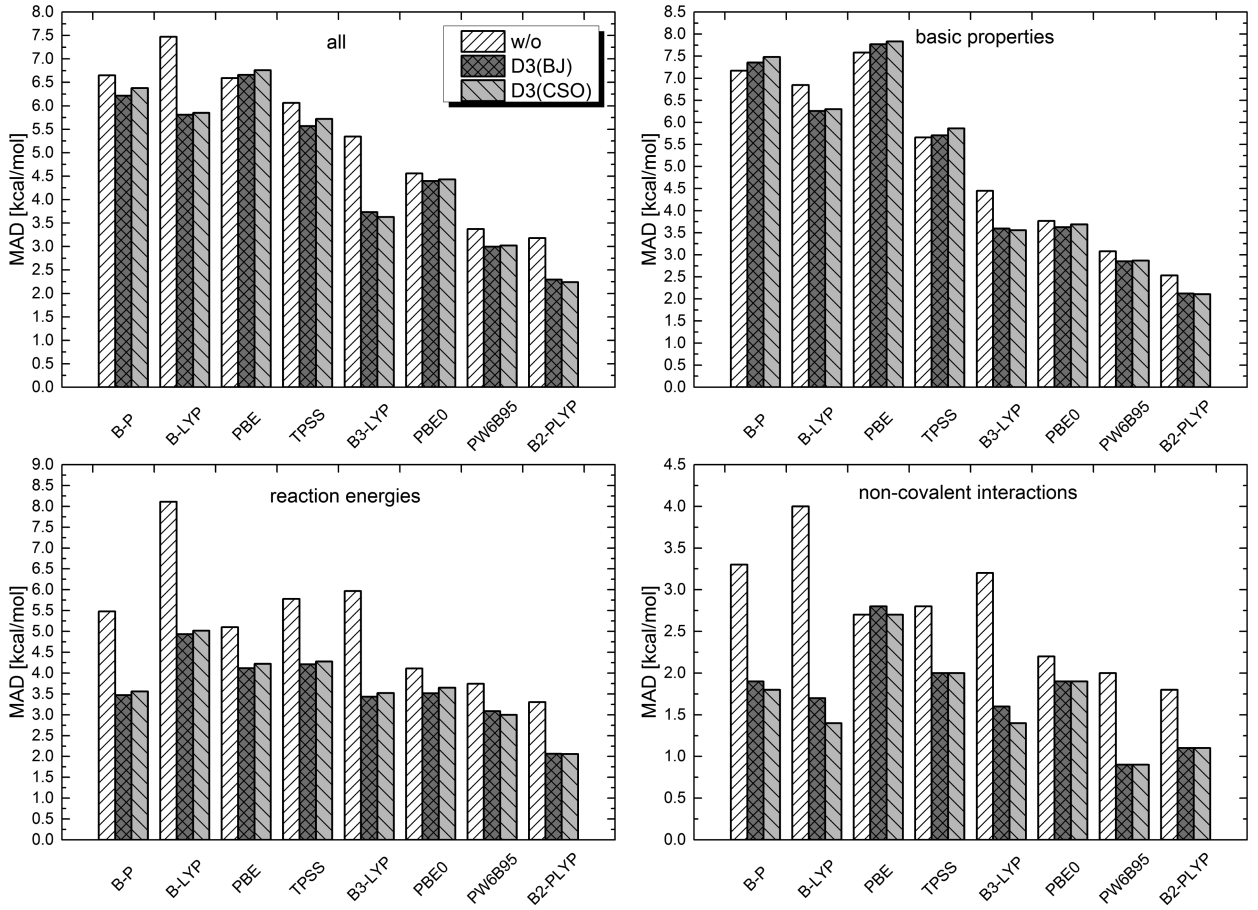


Figure 4. Mean absolute deviations (MAD) of eight different DFAs for the full GMTKN30 set and the three subcategories. Values given without dispersion correction, -D3(BJ) correction and -D3(CSO) correction. Functionals are ordered by categories: GGAs, meta-GGA, exchange-hybrids, exchange-correlation hybrid.

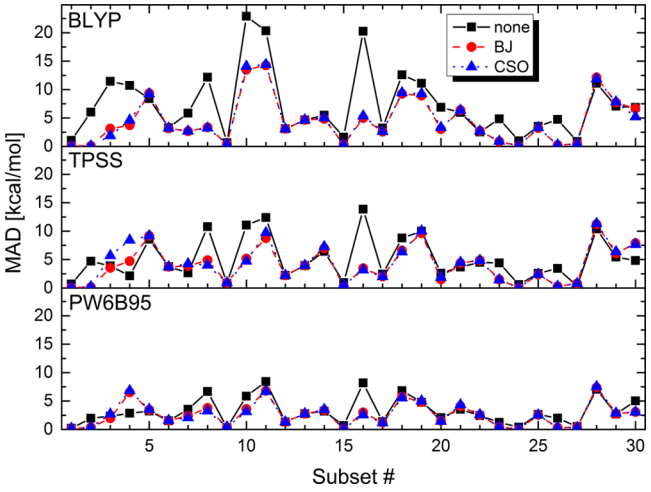


Figure 5. Mean absolute deviations (MAD) of three representative DFAs for individual subsets of the GMTKN30. Values given without dispersion correction, -D3(BJ) correction and -D3(CSO) correction. The subsets are ordered alphabetically.

For subsets SIE11 (#28) and O3ADD6 (#21), dispersion correction leads to slightly worse results for all combinations of functional/correction (see also the Supporting Information). This increases the MAD by up to 1.1 kcal mol^{-1} for the MAD (for B3LYP-D3(CSO) and the SIE11 set). For these subsets, dispersion interactions are not the key interaction. Again, the observed worsening might actually be due to cancellation of error compensation. At least, it should not indicate any fundamental shortcomings of the dispersion correction schemes.

The analysis shows that both dispersion corrections, D3(BJ) and D3(CSO), perform very similarly in this benchmark. This finding also holds for the other DFAs under investigation (see the Supporting Information).

5. CONCLUSION

We have introduced a reformulated version of Grimme's DFT-D3 dispersion correction, which is only dependent on the C_6 coefficient. The role of the higher order correction terms in DFT-D3(BJ) is investigated. It is shown that its purpose is not only to improve interaction energies but also to allow for enough flexibility of the correction to be combinable with standard DFAs. Therefore, it also acts as an interpolation function to vary between the short- and midrange DFA dependent and the long-range DFA independent region.

For the present alternative formulation, a sigmoidal interpolation function is employed instead, while finite damping of Becke and Johnson is kept. For our proposed C_6-only dispersion correction scheme only one parameter needs to be fitted per DFA (instead of three for D3(BJ)). Eight standard DFAs from different classes were parametrized, and their performance for the huge GMTKN30 benchmark set was evaluated against that of D3(BJ). Both corrections performed

DOI: 10.1021/acs.jctc.5b00400
J. Chem. Theory Comput. 2015, 11, 3163−3170

very similarly for the major part of the benchmark set. For the water clusters of the Water27 subset a systematic improvement could be observed, especially for BLYP and B3LYP, though a general improvement for H-bonded complexes could not be found.

Fitting the three empirical parameters of the DFT-D3(BJ) scheme, which are not completely uncoupled and show several local minima, is not straightforward. Because the D3(CSO) dispersion correction consists of only one DFA dependent parameter, it should be much easier to combine with existing or future DFAs. It also has a simpler algebraic form, which might be beneficial for the derivation and computation of its gradient expression. This might be more important in the context of semiempirical approaches, which make use of a dispersion correction,[53,54] or methods like the quantum mechanically derived force field.[55] Note that our approach is not necessarily dependent on the D3 ansatz and other approaches, where C_6 coefficients are obtained differently might as well benefit from the scheme presented here.

In the end, dispersion corrections alone will not remedy the current challenges of DFT,[56] and one should keep in mind the simple and empirical form of the DFT-D scheme. But of course, it works for what it is intended: noncovalent long-range interactions. The present D3(CSO) scheme may not improve quantitatively in comparison to its predecessor, but it reaches the same accuracy with less empirical parameters and also without explicitly resorting to higher than C_6 dispersion coefficient. Against current opinions, the latter clearly puts the significance of higher order terms for empirical dispersion corrections into perspective. But this finding does not challenge the general importance of higher order terms in a strict perturbational series expansion.

■ ASSOCIATED CONTENT

Supporting Information

Results for the fit set and for all GMTKN30 subsets for all presented functionals without correction, with D3(BJ) and with the new D3(CSO) correction as well as an extensive statistically analysis of the performance on the GMTKN30 set for all methods. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00400.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: schwabe@zbh.uni-hamburg.de.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175−180.
(2) Sim, F.; St-Amant, A.; Papai, I.; Salahub, D. R. *J. Am. Chem. Soc.* **1992**, *114*, 4391−4400.
(3) Perdew, J. P.; Schmidt, K. *AIP Conf. Proc.* **2001**, *577*, 1−20.
(4) Klimeš, J.; Michaelides, A. *J. Chem. Phys.* **2012**, *137*, 120901.
(5) Grimme, S. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 211−228.
(6) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
(7) Tkatchenko, A.; DiStasio, R. A., Jr.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012**, *108*, 236402.
(8) *The Theory of Intermolecular Forces*; Stone, A. J., Ed.; Oxford University Press: New York, NY, USA, 2002.
(9) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463−1473.
(10) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(11) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2006**, *28*, 555−569.
(12) Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 6−9.
(13) Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V.; Scoles, G. *J. Chem. Phys.* **2001**, *115*, 8748−8757.
(14) Zimmerli, U.; Parrinello, M.; Koumoutsakos, P. *J. Chem. Phys.* **2004**, *120*, 2693−2699.
(15) Tao, J.; Perdew, J. P.; Ruzsinszky, A. *Phys. Rev. B* **2010**, *81*, 233102.
(16) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(17) Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2006**, *124*, 174104.
(18) Steinmann, S. N.; Corminboeuf, C. *J. Chem. Phys.* **2011**, *134*, 044117.
(19) Ikabata, Y.; Nakai, H. *Int. J. Quantum Chem.* **2015**, 309−324.
(20) Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2014**, *141*, 141101.
(21) Tao, J.; Perdew, J. P.; Ruzsinszky, A. *Proc. Natl. Acad. Sci.* **2012**, *109*, 18−21.
(22) Tao, J.; Fuang, Y.; Hao, P.; Scuseria, G. E.; Ruzsinszky, A.; Perdew, J. *J. Chem. Phys.* **2015**, *142*, 024312.
(23) Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1984**, *80*, 3726−3741.
(24) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515−524.
(25) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(26) Liu, Y.; Goddard, W. A. I. *Mater. Trans.* **2009**, *50*, 1664−1670.
(27) Koide, A. *J. Phys. B: At. Mol. Phys.* **1976**, *9*, 3173−3183.
(28) Hanke, F. *J. Comput. Chem.* **2011**, *32*, 1424−1430.
(29) Starkschall, G. *J. Chem. Phys.* **1972**, *56*, 2801−2806.
(30) Thakkar, A. J.; Hettema, H.; Wormer, P. E. S. *J. Chem. Phys.* **1992**, *97*, 3252−3257.
(31) Goerigk, L.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670−6688.
(32) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 291−309.
(33) Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427−2438.
(34) *TURBOMOLE*, V6.4, 2012, developed by University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007; TURBOMOLE GmbH, since 2007. Available from http://www.turbomole.com (accessed May 26, 2015).
(35) Treutler, O.; Ahlrichs, R. *J. Chem. Phys.* **1995**, *102*, 346−354.
(36) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.
(37) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1995**, *103*, 2975−2988.
(38) Hättig, C. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59−66.
(39) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175−3183.
(40) Schwabe, T. *J. Phys. Chem. A* **2013**, *117*, 2879−2883.
(41) Neese, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73−78.
(42) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(43) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822−8824.
(44) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.
(45) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(46) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(47) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
(48) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(49) Stephens, P.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(50) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(51) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656−5667.
(52) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.
(53) Brandenburg, J. G.; Grimme, S. *J. Phys. Chem. Lett.* **2014**, *5*, 1785−1789.
(54) Korth, M. *J. Chem. Theory Comput.* **2010**, *6*, 3808−3816.

DOI: 10.1021/acs.jctc.5b00400
J. Chem. Theory Comput. 2015, 11, 3163−3170

(55) Grimme, S. *J. Chem. Theory Comput.* **2014**, *10*, 4497−4514.
(56) Cohen, A. J.; Mori-Sáchez, P.; Yang, W. *Chem. Rev.* **2012**, *112*, 289−320.

DOI: 10.1021/acs.jctc.5b00400
J. Chem. Theory Comput. 2015, 11, 3163−3170